AUREL WINTNER

SPEKTRALTHEORIE DER UNENDLICHEN MATRIZEN

A Prefatory Note to this Electronic Republication of

Spektraltheorie der Unendlichen Matrizen

Aurel Wintner

S. Hirzel Verlag, Leipzig 1929

During the period of World War II the above monograph went out of print and was not republished thereafter. As Aurel Wintner's only child and heir, I am the copyright holder. In addition, S. Hirzel Verlag graciously has returned to me all publication rights. I am not able for myself to judge properly the material here presented by my father, but several mathematicians recently have encouraged me to republish the work, on the grounds that it contains some noteworthy results that at times in the past may have been attributed to other workers in the field. In sum, Aurel Wintner's contributions may not always have been acknowledged to the extent they might have been. Thus, given the opportunity offered by our electronic age, it has seemed to me that I simply should allow my father's results to speak for themselves, from the year 1929, to those who can read and understand them.

Claude Wintner
cwintner@haverford.edu
April, 2015

SPEKTRALTHEORIE
DER UNENDLICHEN MATRIZEN

Einführung in den analytischen Apparat der Quantenmechanik

Von

AUREL WINTNER

Mit einer Einleitung

Von

LEON LICHTENSTEIN
o. ö. Professor der Mathematik
an der Universität Leipzig

VERLAG VON S. HIRZEL IN LEIPZIG · 1929

DRUCK VON OSCAR BRANDSTETTER IN LEIPZIG

Vorwort.

Diese Schrift soll eine Einführung in die allgemeine Theorie der linearen Analysis von unendlich vielen Variablen geben. Als ich auf Anregung und Veranlassung von Herrn Professor Lichtenstein die Abfassung des Buches übernommen habe, hoffte ich zwar, bei dieser Gelegenheit auch die nichtlinearen Methoden meines näheren Arbeitsgebietes im Zusammenhang darstellen zu können; leider ließ sich dies nicht durchführen, die Schrift ist ohnehin noch einmal so lang geworden, als es von dem Verlage ursprünglich geplant war. So mußte ich u. a. auch darauf verzichten, auf die Beziehungen zu der Darstellungstheorie, insbesondere zu den neueren Untersuchungen von Weyl über kontinuierliche Gruppen, näher einzugehen, obwohl sich dazu reichliche Gelegenheit geboten hätte.

Die Lehre von den unendlich vielen Veränderlichen ist ein weites, noch nicht lange erschlossenes Gebiet, das in der Analysis und ihren Anwendungen schon heute eine zentrale Stellung einnimmt, in steter Entwicklung begriffen ist und mit seinem kräftigen Apparate überall, sogar in der Zahlentheorie, zu weitgehenden Hoffnungen berechtigt. Die Hilfsmittel, die man dabei benötigt, sind freilich ziemlich heterogener Natur und z. Z. nicht einem jeden geläufig. Da diese Schrift sich dem Wunsche des Verlages entsprechend vor allem an Anfänger wenden sollte, so waren daher manche Kompromisse unumgänglich. So sind z. B. die Kapitel I, II und V absichtlich etwas breit gehalten, während an anderen Stellen, etwa in dem Kapitel IV, an den Leser größere Anforderungen gestellt werden; manchmal war für mich das wesentliche, ihn zum Studium der Originalarbeiten anzuregen.

Da ich kein Handbuch, sondern eine Einführung zu schreiben hatte, glaubte ich meinem Ziele auf diesem Wege am besten nahekommen zu können. Was nicht behandelt wurde, blieb jedenfalls nicht deshalb unbeachtet, weil ich es nicht für wichtig oder interessant hielt. — Was in dem Buch sachlich oder methodisch neu ist, wird der kundige Leser leicht bemerken können.

Als eine wesentliche Erleichterung meiner Aufgabe empfand ich es, daß ich mich von Anfang an der ins Einzelne gehenden Ratschläge von Herrn Professor Lichtenstein erfreuen durfte, der sogar die Korrekturen mitgelesen hat, und daß mir mein lieber Freund, Herr Privatdozent Dr. E. Hölder, bei der Abfassung des Manuskriptes in aufopfernder Weise unterstützend zur Seite stand. Herrn Professor F. Levi verdanke ich manche wertvolle Bemerkungen und freundliche Hilfe beim Lesen der Korrekturfahnen. Einige Bogen hat während seines Leipziger Aufenthaltes auch Herr Dr. W. Nikliborc gelesen. Allen diesen Herren gilt mein bester Dank, vor allem aber dem Verlagshause S. Hirzel, das mich mit der Abfassung dieser Schrift betraut hat und meinen verschiedensten Wünschen vor und während der Drucklegung stets in der liebenswürdigsten Weise entgegengekommen ist.

Ein Sachregister erübrigt sich mit Rücksicht auf die ständigen Rückverweisungen im Texte, sowie auf S. IX—XII.

Leipzig, den 15. Juni 1929.

A. Wintner.

Einleitung.

Durch die neueste an Werner Heisenberg anknüpfende Entwicklung der Quantentheorie hat die von Hilbert begründete, von Hellinger und Toeplitz erfolgreich fortgeführte Theorie der unendlichen Matrizen eine erneute Aktualität gewonnen. Während für die Anwendungen bis dahin in erster Linie die Theorie vollstetiger Formen von Bedeutung war, tritt jetzt gerade die tiefliegende Hilbert-Hellingersche Spektraltheorie beschränkter Formen in den Vordergrund. Die Hermiteschen Formen der Matrizenmechanik sind übrigens nicht einmal beschränkt. Eine lückenlose, mathematisch befriedigende Theorie der quantentheoretischen Matrizen ist zurzeit noch ein Desideratum. An der Weiterentwicklung der Theorie sind, wie wir sehen, Analysis und theoretische Physik in gleichem Maße interessiert.

Dieses Buch ist aus den Vorträgen hervorgegangen, die der Verfasser vor einem kleinen Kreise von Mitgliedern des hiesigen Mathematischen Seminars im Wintersemester 1927/28 gehalten hatte, und versucht, den berechtigten Wünschen der Physiker und Mathematiker gerecht zu werden. Die Aufgabe ist nicht ganz leicht und kaum ohne Zugeständnisse und Opfer zu lösen. So bringt das erste, den algebraischen Grundlagen gewidmete Kapitel dem Mathematiker nur methodisch etwas Neues; der Verfasser behandelt hier endliche Matrizen, — die Beweise sind meist so angelegt, daß sie sich zum Teil unmittelbar auf die unendlichen Matrizen übertragen lassen. Das zweite Kapitel beschäftigt sich mit einigen den Physikern weniger geläufigen fundamentalen Begriffen der reellen Analysis, Funktionen beschränkter Schwankung, Stieltjesschen Integralen, konvergenten Folgen von Funktionen u. dgl. Auch hier wird der Mathematiker noch nicht auf seine Kosten kommen, auch weil mit Rücksicht auf den verfügbaren Raum mancher Beweis nur angedeutet werden konnte. Der Lebesguesche Integralbegriff wird in dem Werk nicht gebraucht, was einen Verzicht auf eine lückenlose Darstellung einiger grundlegenden Ergebnisse der bedauerlicherweise immer noch nicht bequem zugänglichen Hellinger-

schen Dissertation zur Folge hatte. — Trotz aller dieser Zugeständnisse wird der mathematisch weniger gewandte Physiker gelegentlich finden, die Spektraltheorie der unendlichen Matrizen lese sich nicht immer ganz leicht. Dem mag entgegengehalten werden, daß das Buch der Natur nun einmal in mathematischer Sprache geschrieben sei. Seitens des Verfassers ist alles getan, um diese Sprache verständlich zu gestalten, — getan mit viel Geschick, Sachkenntnis und Eleganz.

Die Entwicklungen des dritten und der folgenden Kapitel führen über die eigenartig gefaßte Theorie beschränkter Formen zu gewissen nichtbeschränkten und nicht Hermiteschen Matrizen hinüber. In einigen sehr bemerkenswerten, in der letzten Zeit erschienenen Publikationen ist es dem Verfasser gelungen, für die von ihm so genannten halbbeschränkten und fastperiodischen Matrizen zu überraschend eleganten, abschließenden Ergebnissen zu kommen. Mit den halbbeschränkten Matrizen ist eine mathematisch befriedigende Theorie der quantenmechanischen Matrizen von wasserstoffähnlichem Spektrum gewonnen, während die fastperiodischen Matrizen die klassische Mechanik in die Sprache der unendlichen Matrizen übersetzen. An die fastperiodischen Matrizen schließen sich Betrachtungen an, die an gewisse störungstheoretische Problemstellungen und Ergebnisse von Lagrange und Levi-Civita anknüpfen. An diesen Dingen und manchem anderen in dem Buch, so an der Spektraltheorie der unitären und allgemeiner der normalen beschränkten Matrizen wird der Mathematiker seine helle Freude haben, die gleiche Freude, die ich empfand, als ich dieses Erstlingswerk des jungen, hochbegabten Verfassers las.

Leon Lichtenstein.

Inhaltsverzeichnis.

Erstes Kapitel.

Algebraische und formale Grundlagen.

Zweites Kapitel.
Analytische Hilfsmittel.

Drittes Kapitel.
Die beschränkten unendlichen Matrizen.

Viertes Kapitel.
Theorie der Spektralmatrix.

Erstes Kapitel.

Algebraische und formale Grundlagen.

§ 1. Methodische Vorbemerkung.

Es sei zu jedem der $n^2 (\geqq 1)$ Paare (i, k) von ganzen, der Bedingung $1 \leqq i \leqq n,\ 1 \leqq k \leqq n$ genügenden Zahlen i, k eine reelle oder komplexe Größe a_{ik} zugeordnet. Es handelt sich also um ein geordnetes System

$$(1) \qquad \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{Vmatrix}$$

von n^2 Elementen, das eine Matrix genannt wird und kurz mit $\| a_{ik} \|_{(n)}$, oder, wenn kein Mißverständnis zu befürchten ist, mit $\| a_{ik} \|$ bezeichnet werden möge. Die Zahlen $a_{i1}, a_{i2}, \ldots, a_{in}$; $a_{1k}, a_{2k}, \ldots, a_{nk}$; $a_{11}, a_{22}, \ldots, a_{nn}$ bilden bzw. „die i-te Zeile“, „die k-te Spalte (Kolonne)“ und „die Diagonale“ von $\| a_{ik} \|$. Man spricht von Matrizen gelegentlich auch dann, wenn die Anzahl der Zeilen mit derjenigen der Spalten nicht übereinstimmt. Doch werden wir es mit diesen nur rechteckigen und nicht zugleich quadratischen Schemata nicht zu tun haben. — Die Zahl n heißt der Grad von $\| a_{ik} \|$. Man sagt auch, $\| a_{ik} \|$ sei eine n-äre (binäre, ternäre usf.) Matrix; die 1-ären Matrizen sind die gewöhnlichen, reellen oder komplexen Zahlen.

Unser eigentliches Ziel ist die Behandlung von unendlichen Matrizen $(n = +\infty)$. Doch müssen wir uns vorher mit den endlichen Matrizen ausführlich beschäftigen, zum Teil zwecks Orientierung in den algebraisch-formalen Grundlagen der transzendenten Probleme, zum Teil aber zur Erleichterung des Grenzüberganges $n \to +\infty$, der unter Zugrundelegung der im Kap. II zu entwickelnden funktionentheoretischen Hilfsmittel erfolgen wird. Bei den unendlichen Matrizen wird es sich vor allem um Hermitesche und unitäre Matrizen handeln. Beliebige Matrizen werde ich des näheren nur im Falle $n < +\infty$ behandeln. Doch sind die dabei ver-

wandten Methoden zum Teil auch bei unendlichen Matrizen valent. — Alle unsere Betrachtungen über endliche Matrizen sind in diesem Sinne präpariert, also sogleich dem Falle $n = +\infty$ angepaßt. So werden wir Terminologie und Beweisführung nicht immer auf die scheinbar am nächsten liegende Art und Weise, sondern derart zu führen haben, daß sie mutatis mutandis (oder sogar ohne weiteres) auch im Falle unendlicher Matrizen in Kraft bleiben. — Die scheinbaren Umwege erweisen sich also als methodische Notwendigkeiten: auf unendliche Matrizen können nur „primitive" Schlußweisen übertragen werden, und Sätze, die solchen Schlußweisen nicht zugänglich sind, pflegen im Falle $n = +\infty$ geradezu falsch zu sein. Unsere „Umwege" werden sogar die Vermeidung mancher Wiederholungen erlauben. Man gewinnt nur auf diese Weise einen näheren Einblick in die wahre Natur der Probleme.

Der Grad soll in diesem Kapitel beliebig, jedoch immer derselbe sein, sofern nichts anderes vermerkt wird.

§ 2. Matrizenpolynome.

Die Matrix ist durch die Vorgabe von n^2 Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_j, \ldots, \alpha_{n^2}$ noch nicht festgelegt. Man muß vielmehr auch den Ort von α_j in der Tabelle angeben, d. h. man muß wissen, welchem Wertepaare (i, k) der Zeiger j entspricht. Dementsprechend heißen die beiden Matrizen $\|a_{ik}\|$, $\|b_{ik}\|$ dann und nur dann *gleich*, $\|a_{ik}\| = \|b_{ik}\|$, wenn die *entsprechenden* Elemente identisch sind, so daß also eine Matrizengleichung nur eine Abkürzung für ein System von n^2 gewöhnlichen Gleichungen $a_{ik} = b_{ik}$ darstellt. Die Matrix $\|a_{ik}\|$ kann dabei als *eine* hyperkomplexe Zahl mit n^2 Einheiten aufgefaßt werden, so daß wir sie mit einem einzigen Buchstaben bezeichnen können. Zu diesem Zwecke soll der entsprechende große deutsche Buchstabe verwendet werden, so daß wir z. B. $\|a_{ik}\| = \mathfrak{A}$ setzen.

Sind A und B gewöhnliche Zahlen, so verstehen wir unter $A\mathfrak{A} + B\mathfrak{B}$ die Matrix $\|A a_{ik} + B b_{ik}\|$. Wir setzen also $\|a_{ik}\| \pm \|b_{ik}\| = \|a_{ik} \pm b_{ik}\|$. Es sei $A\|a_{ik}\| = \|A a_{ik}\|$. Da die gewöhnlichen Zahlen sich der Addition gegenüber kommutativ und assoziativ verhalten, so gilt dies auch von den Matrizen, so daß also $\mathfrak{A} + \mathfrak{B} = \mathfrak{B} + \mathfrak{A}$, $(\mathfrak{A} + \mathfrak{B}) + \mathfrak{C} = \mathfrak{A} + (\mathfrak{B} + \mathfrak{C})$ ist. Es kann daher dem Symbol $\mathfrak{A} + \mathfrak{B} + \mathfrak{C}$ ein von der Reihenfolge und von der Gruppierung der Summanden unabhängiger Sinn erteilt werden, etwa $(\mathfrak{A} + \mathfrak{B}) + \mathfrak{C}$. Sind ${}^{(1)}\mathfrak{A}$, ${}^{(2)}\mathfrak{A}, \ldots$ beliebige Matrizen (alle vom selben Grade n), so setzen wir

$$(2) \qquad \sum_{\nu=1}^{N} \|{}^{(\nu)}a_{ik}\| = \|\sum_{\nu=1}^{N} {}^{(\nu)}a_{ik}\|,$$

und zwar auch im Falle $N = +\infty$, sofern alle n^2 Reihen rechterhand konvergent sind. Wir werden es nur mit Fällen zu tun haben, wo jede dieser Reihen absolut konvergiert.

Das Produkt $\mathfrak{A}\mathfrak{B}$ zweier Matrizen vom selben Grad wird zu $\|\sum_{j=1}^{n} a_{ij} b_{jk}\|$ festgelegt, es ist wieder eine Matrix vom Grade n. Nach Definition ist also

$$\|a_{ik}\|\,\|b_{ik}\| = \|c_{ik}\|, \qquad c_{ik} = a_{i1} b_{1k} + a_{i2} b_{2k} + \cdots + a_{in} b_{nk}.$$

Hingegen ist

$$\|b_{ik}\|\,\|a_{ik}\| = \|d_{ik}\|, \qquad d_{ik} = b_{i1} a_{1k} + b_{i2} a_{2k} + \cdots + b_{in} a_{nk}.$$

Da die n^2 Gleichungen $c_{ik} = d_{ik}$ bei beliebigen Matrizenpaaren $\mathfrak{A}, \mathfrak{B}$ nicht erfüllt sind, so ist die Richtigkeit der Beziehung $\mathfrak{A}\mathfrak{B} = \mathfrak{B}\mathfrak{A}$ eine besondere Eigenschaft des Matrizenpaares, die ihm zukommt oder nicht zukommt. Zwei Matrizen sind also im allgemeinen nicht *vertauschbar*, die Matrizenmultiplikation ist keine kommutative Operation. Hingegen ist sie, wie man leicht nachrechnet, gewiß distributiv in bezug auf Matrizenaddition, d. h. es gilt bei drei beliebigen Matrizen die Beziehung $(\mathfrak{A} + \mathfrak{B})\mathfrak{C} = \mathfrak{A}\mathfrak{C} + \mathfrak{B}\mathfrak{C}$. Endlich ist die Matrizenmultiplikation assoziativ: man hat $(\mathfrak{A}\mathfrak{B})\mathfrak{C} = \mathfrak{A}(\mathfrak{B}\mathfrak{C})$. Man bilde nämlich zunächst die Matrix $\mathfrak{A}\mathfrak{B}$ und multipliziere sie von hinten mit $\mathfrak{C}$. Man bilde sodann die Matrix $\mathfrak{B}\mathfrak{C}$ und multipliziere sie von vorn mit $\mathfrak{A}$. Beidemal ergibt sich dieselbe Matrix, nämlich

$$\|g_{ik}\| = \|\sum_{j=1}^{n}\sum_{l=1}^{n} a_{ij} b_{jl} c_{lk}\|,$$

die also — ohne Klammern — einfach mit $\mathfrak{A}\mathfrak{B}\mathfrak{C}$ bezeichnet werden kann. — Sind die m Matrizen ${}^{(\nu)}\mathfrak{A}$ paarweise vertauschbar und bedeutet $(p_1, p_2, \ldots, p_m)$ eine beliebige Permutation von $(1, 2, \ldots, m)$, so ist $\mathfrak{A}_1 \mathfrak{A}_2 \ldots \mathfrak{A}_m = \mathfrak{A}_{p_1} \mathfrak{A}_{p_2} \ldots \mathfrak{A}_{p_m}$, da bekanntlich jede Permutation von m Elementen durch endlich häufige Wiederholung von Vertauschungen je zweier benachbarten Elemente gewonnen werden kann. Da mit sich selbst jede Matrix $\mathfrak{A}$ vertauschbar ist, können daher die Potenzen $\mathfrak{A}^m$; $m = 2, 3, \ldots$, eindeutig erklärt werden. Wir setzen $\mathfrak{A}^1 = \mathfrak{A}$ und $\mathfrak{A}^0 = \mathfrak{E}$, unter $\mathfrak{E}$ die „Einheitsmatrix"

$$\mathfrak{E} = \|\delta_{ik}\|; \qquad \delta_{ii} = 1, \quad \delta_{ik} = 0 \quad \text{für} \quad i \neq k$$

verstanden. Für jede Matrix $\mathfrak{A}$ gilt, wie man leicht ausrechnet,

$$\mathfrak{A}\mathfrak{E} = \mathfrak{E}\mathfrak{A} = \mathfrak{A}, \quad \text{ferner} \quad \mathfrak{A}\mathfrak{D} = \mathfrak{D}\mathfrak{A} = \mathfrak{D}, \quad \mathfrak{A} \pm \mathfrak{D} = \mathfrak{A},$$

wenn $\mathfrak{O}$ die „Nullmatrix“ $\|0\|$ bezeichnet, deren alle Elemente verschwinden. Hierbei ist zu beachten, daß, wie das Beispiel des binären Matrizenpaares

$$(3)\qquad \mathfrak{A} = \begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix}, \quad \mathfrak{B} = \begin{Vmatrix} 0 & 0 \\ 0 & 1 \end{Vmatrix}$$

zeigt, $\mathfrak{A}\mathfrak{B} = \mathfrak{O}$ auch dann möglich ist, wenn sowohl $\mathfrak{A}$ als auch $\mathfrak{B} \neq \mathfrak{O}$ ausfällt (ebenso wie das Produkt zweier Vektoren von positivem Betrag verschwinden kann).

Ist $F(z) = \sum\limits_{\nu=0}^{q} A_\nu z^\nu$ eine ganze rationale Funktion, so verstehen wir unter $F(\mathfrak{A})$ die Matrix

$$F(\mathfrak{A}) = \sum_{\nu=0}^{q} A_\nu \mathfrak{A}^\nu = A_0 \mathfrak{E} + A_1 \mathfrak{A} + A_2 \mathfrak{A}^2 + \cdots + A_q \mathfrak{A}^q.$$

— Ist $G(z, w) = \sum\sum A_{\nu\mu} z^\nu w^\mu$ eine ganze rationale Funktion zweier Veränderlichen, so gehört ebenso zu dem Symbol $G(\mathfrak{A}, \mathfrak{B})$ eine und nur eine Matrix, wenn $\mathfrak{A}$ und $\mathfrak{B}$ vertauschbar sind. Sind sie aber nicht vertauschbar, so ist $G(\mathfrak{A}, \mathfrak{B})$ im allgemeinen mehrdeutig: es genügt nicht, die *Funktion* $G(z, w)$ anzugeben. Ist z. B. $A_{22} = 1$ und sonst $A_{\nu\mu} = 0$, so ist $G(z, w) = z^2 w^2 = zw\,zw = wzzw$ usf., während im allgemeinen $\mathfrak{A}^2\mathfrak{B}^2 \neq \mathfrak{A}\mathfrak{B}\mathfrak{A}\mathfrak{B} \neq \mathfrak{B}\mathfrak{A}\mathfrak{A}\mathfrak{B} \neq \ldots$ ausfällt. Da hierbei an und für sich alle Permutationen gleichberechtigt sind, so muß man, um $G(\mathfrak{A}, \mathfrak{B})$ für alle Fälle eindeutig festzulegen, außer der Funktion $G(z, w)$ die Gruppierung der Faktoren in jedem Glied des näheren angeben, sofern $G(z, w)$ Glieder enthält, in denen beide Veränderlichen eintreten.

§ 3. Diagonalmatrix. Begleitende Matrix. Unitäre Matrix.

Eine Matrix $\mathfrak{H} = \|h_{ik}\|$, in der $h_{ik} = h_{ik}\delta_{ik}$ ausfällt, so daß die außerhalb der Diagonale stehenden Elemente verschwinden, nennt man eine Diagonalmatrix. Setzt man $h_{ii} = h_i$, so ist $h_{ik} = h_{ik}\delta_{ik} = h_{ii}\delta_{ik}$, also $\mathfrak{H} = \|h_i\delta_{ik}\|$. Man überzeugt sich leicht, daß $\|h_i\delta_{ik}\|\,\|g_i\delta_{ik}\| = \|g_i h_i \delta_{ik}\|$ gilt, woraus wegen $h_i g_i = g_i h_i$ folgt, daß zwei Diagonalmatrizen stets vertauschbar sind.

Jeder Matrix $\mathfrak{A} = \|a_{ik}\|$ ordnen wir die drei Matrizen $\overline{\mathfrak{A}} = \|\bar{a}_{ik}\|$, $\mathfrak{A}' = \|a_{ki}\|$, $\mathfrak{A}^* = \|\bar{a}_{ki}\|$ zu. Hierbei ist unter $\|a_{ki}\|$ die Matrix verstanden, die aus $\|a_{ik}\|$ durch Ersatz der i-ten Zeile (Spalte) mit der i-ten Spalte (Zeile) entsteht, und $\bar{A}$ bezeichnet die zu A konjugiert

komplexe Größe, so daß

(4) $A\bar{A} = |A|^2, \quad \overline{A \pm B} = \bar{A} \pm \bar{B}, \quad \overline{AB} = \bar{A}\,\bar{B}, \quad \overline{\left(\frac{A}{B}\right)} = \frac{\bar{A}}{\bar{B}}$

gilt und $A = \bar{A}$ die Realität von A ausspricht. $\bar{\mathfrak{A}}$ heißt die konjugierte, $\mathfrak{A}'$ die transponierte, $\mathfrak{A}^*$ die begleitende Matrix von $\mathfrak{A}$. Offenbar ist $\bar{\bar{\mathfrak{A}}} = \mathfrak{A}$, $\mathfrak{A}'' = \mathfrak{A}$, $\mathfrak{A}^{**} = \mathfrak{A}$ und $\mathfrak{A}^* = \overline{\mathfrak{A}'} = \bar{\mathfrak{A}}'$. Aus $\|a_{ik}\|\,\|b_{ik}\|$ $= \|\sum_{j=1}^{n} a_{ij} b_{jk}\|$ folgt wegen (4) unmittelbar $\overline{(\mathfrak{A}\mathfrak{B})} = \bar{\mathfrak{A}}\,\bar{\mathfrak{B}}$. Hingegen ist $(\mathfrak{A}\mathfrak{B})' = \mathfrak{B}'\mathfrak{A}'$ und daher $(\mathfrak{A}\mathfrak{B})^* = \mathfrak{B}^*\mathfrak{A}^*$. Man hat nämlich

$$(\mathfrak{A}\mathfrak{B})' = \|\sum_{j=1}^{n} a_{ij} b_{jk}\|' = \|\sum_{j=1}^{n} a_{kj} b_{ji}\| = \|b_{ki}\|\,\|a_{ki}\| = \mathfrak{B}'\mathfrak{A}'.$$

Eine Matrix $\mathfrak{U} = \|u_{ik}\|$, bei welcher $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E} =$ Einheitsmatrix ist, heißt unitär. Die Definitionsgleichung geht in der gewöhnlichen Schreibweise in $\sum_{j=1}^{n} u_{ij}\bar{u}_{kj} = \delta_{ik}$ über. Ist also $\mathfrak{U}$ reell, $\bar{u}_{kj} = u_{kj}$, so haben wir es mit den Matrizen zu tun, welche die Richtungskosinus der durch die reellen Drehungen und Spiegelungen im n-dimensionalen (reellen) euklidischen Raume transformierten Achsen darstellen. — Wir werden auf S. 7 sehen, daß zwei (endliche) Matrizen $\mathfrak{C}$, $\mathfrak{D}$, die miteinander in der Beziehung $\mathfrak{C}\mathfrak{D} = \mathfrak{E}$ stehen, gewiß vertauschbar sind. Aus $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$ folgt daher $\mathfrak{U}^*\mathfrak{U} = \mathfrak{E}$ oder $\mathfrak{U}^*\mathfrak{U}^{**} = \mathfrak{E}$, so daß also auch $\mathfrak{U}^*$ unitär ist. Man hat $\mathfrak{E} = \bar{\mathfrak{E}} = \mathfrak{E}^* = \mathfrak{E}'$, aus $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$ folgt daher einerseits $\overline{\mathfrak{U}}\,\overline{\mathfrak{U}^*} = \bar{\mathfrak{U}}\,\bar{\mathfrak{U}}^* = \mathfrak{E}$, andererseits $\mathfrak{U}^{*\prime}\mathfrak{U}' = \mathfrak{E}$, also $\mathfrak{U}'\mathfrak{U}^{*\prime} = \mathfrak{E}$, d. h. $\mathfrak{U}'\mathfrak{U}'^* = \mathfrak{E}$, so daß auch $\bar{\mathfrak{U}}$ und $\mathfrak{U}'$ unitär sind.

§ 4. Die Reziproke.

Die Matrix $\mathfrak{A} = \|a_{ik}\|$ möge nicht entartet genannt werden, wenn das Gleichungssystem

(5) $$\sum_{j=1}^{n} a_{ij} x_j = c_i; \qquad i = 1, 2, \ldots, n$$

bei *allen* Wertsystemen $\{c_i\}$ mindestens eine Lösung zuläßt. Ist dies der Fall, so ersieht man leicht (indem man zunächst $c_1 = 1$ setzt und die übrigen $c_i = 0$, usf.; vgl. übrigens weiter unten), daß es eine Lösung gibt, die sich aus den c_i homogen und linear aufbaut,

(6) $$x_j = \sum_{k=1}^{n} b_{jk} c_k; \qquad j = 1, 2, \ldots, n,$$

wobei die Matrix $\mathfrak{B} = \|b_{ik}\|$ unabhängig von den c_i ist. — Durch Einsetzen folgt

$$(7) \qquad \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{n} b_{jk} c_k = c_i;$$

setzt man hierin $c_i = \delta_{im}$, wobei m irgendwie fest gewählt ist, so bleibt

$$(8) \qquad \sum_{j=1}^{n} a_{ij} b_{jm} = \delta_{im}; \qquad i, m = 1, 2, \ldots, n,$$

d. h. $\mathfrak{A}\mathfrak{B} = \mathfrak{E}$. Es sei jetzt umgekehrt (8) erfüllt. Multipliziert man (8) mit c_m und summiert man über m, so ergibt sich (7), und wird in (7) wieder (6) eingeführt, so erhält man (5). Mithin ist $\mathfrak{A}$ dann und nur dann nicht entartet, wenn es mindestens eine Matrix $\mathfrak{B}$ gibt derart, daß $\mathfrak{A}\mathfrak{B} = \mathfrak{E}$ ausfällt, oder, wie wir sagen wollen, wenn $\mathfrak{A}$ mindestens eine hintere Reziproke hat. Entsprechend heißt $\mathfrak{B}$ vordere Reziproke von $\mathfrak{A}$, wenn $\mathfrak{B}\mathfrak{A} = \mathfrak{E}$ gilt. Würde $\mathfrak{A}$ zwei verschiedene hintere Reziproken ${}^{(1)}\mathfrak{B}_1$, ${}^{(2)}\mathfrak{B}$ haben, so würde (5) bei *jedem* c_i-System sowohl durch ${}^{(1)}x_j = \sum_{k=1}^{n} {}^{(1)}b_{jk} c_k$ als auch durch ${}^{(2)}x_j = \sum_{k=1}^{n} {}^{(2)}b_{jk} c_k$ befriedigt sein. Da mindestens ein ${}^{(1)}b_{jk} \neq {}^{(2)}b_{jk}$ sein soll, so gibt es dann ein Wertsystem $\{c_i\}$ und dazu ein $j = j_0$ derart, daß ${}^{(1)}x_{j_0} \neq {}^{(2)}x_{j_0}$ ausfällt. Dies ist aber nicht möglich. Es ist nämlich aus den Elementen bekannt, daß, wenn (5) bei *jedem* $\{c_i\}$-System lösbar ist [d. h. wenn die Determinante von $\|a_{ik}\|$ nicht verschwindet], die Lösung bei *jedem* $\{c_i\}$-System eindeutig ist (wir werden sehen, daß für unendliche Systeme der entsprechende Satz unrichtig ist). Es sind also nur zwei Fälle möglich. Entweder ist $\mathfrak{A}$ entartet[1]) und es gibt dann keine hintere Reziproke. Oder $\mathfrak{A}$ ist nicht entartet und es gibt dann eine, aber auch nur eine hintere Reziproke. Daraus folgt, daß $\mathfrak{A}$ entweder keine oder *eine* vordere Reziproke hat. Denn ist ${}^{(1)}\mathfrak{B}\mathfrak{A} = \mathfrak{E}$ und ${}^{(2)}\mathfrak{B}\mathfrak{A} = \mathfrak{E}$, so gilt $\mathfrak{A}'\,{}^{(1)}\mathfrak{B}' = \mathfrak{E}$ und $\mathfrak{A}'\,{}^{(2)}\mathfrak{B}' = \mathfrak{E}$, also, da $\mathfrak{A}'$ nur eine hintere Reziproke haben kann, ${}^{(1)}\mathfrak{B}' = {}^{(2)}\mathfrak{B}'$ und daher auch ${}^{(1)}\mathfrak{B} = {}^{(2)}\mathfrak{B}$. — Hat $\mathfrak{A}$ eine hintere Reziproke $\mathfrak{B}$, so ist $\mathfrak{A}\mathfrak{B} = \mathfrak{E}$, folglich $\mathfrak{A}\mathfrak{B}\mathfrak{A} = \mathfrak{A}$, mithin $\mathfrak{A}(\mathfrak{B}\mathfrak{A} - \mathfrak{E}) = \mathfrak{O}$, also $\mathfrak{A}\mathfrak{B} + \mathfrak{A}(\mathfrak{B}\mathfrak{A} - \mathfrak{E}) = \mathfrak{E} + \mathfrak{O}$ und daher $\mathfrak{A}(\mathfrak{B} + \mathfrak{B}\mathfrak{A} - \mathfrak{E}) = \mathfrak{E}$. Da jedoch $\mathfrak{A}$ nur eine hintere Reziproke hat, so muß $\mathfrak{B} + \mathfrak{B}\mathfrak{A} - \mathfrak{E} = \mathfrak{B}$, also $\mathfrak{B}\mathfrak{A} = \mathfrak{E}$ sein, so daß $\mathfrak{B}$ auch vordere Reziproke von $\mathfrak{A}$ ist. Ebenso zeigt man, daß eine vordere Reziproke der Matrix $\mathfrak{A}$ zugleich auch hintere Reziproke von ihr sein muß. — Bei den endlichen Matrizen sind also nur zwei Fälle möglich: entweder hat $\mathfrak{A}$ überhaupt keine Reziproke, oder es gibt eine und nur eine

[1]) Entartet = nicht nichtentartet; nichtentartet = nicht entartet.

Reziproke, die dann hintere und zugleich vordere Reziproke ist. Letzteres ist dann und nur dann der Fall, wenn das Gleichungssystem (5) bei *jedem* $\{c_i\}$-System lösbar ist, d. h. wenn $\mathfrak{A}$ nicht entartet ist oder wenn $\det \mathfrak{A} \neq 0$ gilt.

Die Reziproke von $\mathfrak{A}$ möge, wenn sie existiert, mit $\mathfrak{A}^{-1}$ bezeichnet werden. Bekanntlich ist

$$(9) \qquad \|a_{ik}\|^{-1} = \mathfrak{A}^{-1} = \left\| \frac{(-1)^{i+k}}{\det \mathfrak{A}} \frac{\partial \det \mathfrak{A}}{\partial a_{ik}} \right\|.$$

Da $\mathfrak{A}^{-1}$ vordere *und* hintere Reziproke ist, so ist $\mathfrak{A}$ mit $\mathfrak{A}^{-1}$ vertauschbar.

Es ist trivial, daß $\mathfrak{A}$ und $\overline{\mathfrak{A}}$ nur zugleich entartet sein können: man gehe nur in (5) beiderseits zu den konjugierten Größen über. Ferner ist $\mathfrak{A}$ auch mit $\mathfrak{A}^*$ zugleich entartet. Es genügt zu zeigen, daß $\mathfrak{A}$ und $\mathfrak{A}'$ zugleich entartet sind (der entsprechende Satz ist für unendliche Matrizen, wie wir sehen werden, im allgemeinen unrichtig). Hat nämlich $\mathfrak{A}$ eine Reziproke $\mathfrak{B}$, so gilt, wegen $\mathfrak{A}\mathfrak{B} = \mathfrak{E}$ und da $(\mathfrak{A}\mathfrak{B})' = \mathfrak{B}'\mathfrak{A}'$ ist, auch $\mathfrak{B}'\mathfrak{A}' = \mathfrak{E}' = \mathfrak{E}$, d. h. $\mathfrak{A}'$ hat eine Reziproke, die übrigens die transponierte Matrix der Reziproke von $\mathfrak{A}$ ist, $(\mathfrak{A}')^{-1} = (\mathfrak{A}^{-1})'$. Ebenso folgt $(\mathfrak{A}^*)^{-1} = (\mathfrak{A}^{-1})^*$ und $(\overline{\mathfrak{A}})^{-1} = \overline{(\mathfrak{A}^{-1})}$.

Vertauscht man in einer Determinante die Zeilen mit den Spalten, so ändert sich der Wert der Determinante nicht. Es ist also $\det \mathfrak{A}' = \det \mathfrak{A}$. Wegen (4) gilt ferner $\det \overline{\mathfrak{A}} = \overline{\det \mathfrak{A}}$, also auch $\det \mathfrak{A}^* = \overline{\det \mathfrak{A}}$. Der Multiplikationssatz der Determinanten besagt, daß $\det(\mathfrak{A}\mathfrak{B}) = (\det \mathfrak{A})(\det \mathfrak{B})$ ist. Da die Multiplikation gewöhnlicher Zahlen kommutativ ist, so folgt daraus $\det(\mathfrak{A}\mathfrak{B}) = \det(\mathfrak{B}\mathfrak{A})$ auch für den Fall $\mathfrak{A}\mathfrak{B} \neq \mathfrak{B}\mathfrak{A}$. Wegen $\mathfrak{A}\mathfrak{A}^{-1} = \mathfrak{E}$ und $\det \mathfrak{E} = 1$ ist $(\det \mathfrak{A})(\det \mathfrak{A}^{-1}) = 1$. Es ist also $\det(\mathfrak{A}^n) = (\det \mathfrak{A})^n$; $n = 0, 1, 2, \ldots$, und dies gilt, wenn $\mathfrak{A}^{-1}$ vorhanden ist, auch für $n = -1, -2, -3, \ldots$. Ist $\mathfrak{A}^{-1}$ vorhanden, so ist auch $\mathfrak{A}^n$ nicht entartet und $(\mathfrak{A}^n)^{-1} = (\mathfrak{A}^{-1})^n$ für $n = 1, 2, \ldots$ (für $n = 0$ ist dies wegen $\mathfrak{B}^0 = \mathfrak{E} = \mathfrak{E}^n$; $n = 0, \pm 1, \pm 2, \ldots$ trivial). Sind nämlich zwei Matrizen $\mathfrak{C}$, $\mathfrak{D}$ nicht entartet, so ist es auch das Produkt $\mathfrak{C}\mathfrak{D}$ nicht, und wegen

$$(\mathfrak{C}\mathfrak{D})(\mathfrak{D}^{-1}\mathfrak{C}^{-1}) = \mathfrak{C}(\mathfrak{D}\mathfrak{D}^{-1})\mathfrak{C}^{-1} = \mathfrak{C}\mathfrak{E}\mathfrak{C}^{-1} = \mathfrak{C}\mathfrak{C}^{-1} = \mathfrak{E}$$

gilt $(\mathfrak{C}\mathfrak{D})^{-1} = \mathfrak{D}^{-1}\mathfrak{C}^{-1}$.

Wir wollen zum Schluß die aus der Determinantentheorie bekannten Sätze über Systeme von n linearen Gleichungen mit n Unbekannten auf eine determinantenfreie Weise zusammenstellen:

Wenn die Gleichungen

$$(10)\qquad \sum_{k=1}^{n} a_{ik} x_k = c_i, \qquad i = 1, 2, \ldots, n$$

bei jedem c_i-System mindestens eine Lösung haben, oder wenn sie bei mindestens einem c_i-System eine und nur eine Lösung zulassen, so haben sowohl sie als auch die transponierten Gleichungen

$$(11)\qquad \sum_{k=1}^{n} a_{ki} x_k = c_i, \qquad i = 1, 2, \ldots, n$$

bei jedem c_i-System eine und nur eine Lösung. Sonst haben die zu $c_1 = c_2 = \cdots = c_n = 0$ gehörigen, homogenen Gleichungen (10) außer der trivialen Lösung $x_1 = x_2 = \cdots = x_n = 0$ noch weitere Lösungen (Alternativsatz). Letzteres ist dann und nur dann der Fall, wenn die Matrix $\| a_{ik} \|$ die folgende Eigenschaft hat: es gibt Multiplikatoren $\mu_1, \mu_2, \ldots, \mu_n$, die nicht alle verschwinden und trotzdem derart sind, daß

$$(12)\qquad \sum_{i=1}^{n} \mu_i \sum_{k=1}^{n} a_{ik} x_k \equiv 0$$

ausfällt. Die inhomogenen Gleichungen sind dabei dann und nur dann lösbar (freilich nicht *eindeutig* lösbar), wenn die c_i derart sind, daß

$$(13)\qquad \sum_{i=1}^{n} \mu_i c_i = 0$$

gilt, wobei die μ_i irgendwelche, der Bedingung (12) genügende Multiplikatoren bezeichnen. Die zu $\mathfrak{A}$ gehörigen homogenen Gleichungen haben dabei nur eine endliche Anzahl ($\geqq 0$) linear unabhängiger Lösungen (vgl. § 5). Diese Anzahl ist (wenn sie möglichst groß gewählt wird) unabhängig von der Wahl der zugrunde gelegten linear unabhängigen Lösungen. Und zwar ist diese Anzahl dieselbe für $\mathfrak{A}$ und $\mathfrak{A}'$.

§ 5. Vektoren.

Ein in einer bestimmten Reihenfolge gegebenes System $\mathfrak{a} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ von n reellen oder komplexen Zahlen α_i heißt ein (n-dimensionaler[1])) Vektor und α_i die i-te Komponente von $\mathfrak{a}$. Ein anderer Vektor, $\mathfrak{b} = (\beta_1, \beta_2, \ldots, \beta_n)$, ist gleich $\mathfrak{a}$, wenn $\alpha_i = \beta_i$; $i = 1, 2, \ldots, n$ gilt. Der Vektor $\mathfrak{o} = (0, 0, \ldots, 0)$ heißt Nullvektor und der Vektor $\mathfrak{e}_i = (\delta_{1i}, \delta_{2i}, \ldots, \delta_{ni})$, dessen i-te Komponente $= 1$ ist, während die

[1]) Der n-dimensionale Vektor hängt von den $2n$ *reellen* Parametern ab.

übrigen $n-1$ Komponenten verschwinden, der i-te Einzelvektor. Die Addition von Vektoren, sowie die Multiplikation eines Vektors mit gewöhnlichen Zahlen, endlich der konjugierte Vektor eines Vektors werden durch die Definitionen

$$\mathfrak{a} \pm \mathfrak{b} = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n); \quad c\,\mathfrak{a} = \mathfrak{a}\,c = (c\,\alpha_1, c\,\alpha_2, \ldots, c\,\alpha_n);$$
$$\bar{\mathfrak{a}} = (\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_n)$$

eindeutig festgelegt. Als das Produkt zweier Vektoren wird die gewöhnliche Zahl

$$\mathfrak{a}\mathfrak{b} = \mathfrak{b}\mathfrak{a} = \alpha_1\beta_1 + \alpha_2\beta_2 + \cdots + \alpha_n\beta_n$$

bezeichnet. Ist $\mathfrak{a}\bar{\mathfrak{b}} = 0$, oder, was offenbar dasselbe ist, $\bar{\mathfrak{a}}\mathfrak{b} = 0$, so sagt man, daß die beiden Vektoren $\mathfrak{a}, \mathfrak{b}$ zueinander orthogonal sind. Das Produkt $\mathfrak{a}\bar{\mathfrak{a}} = \alpha_1\bar{\alpha}_1 + \cdots + \alpha_n\bar{\alpha}_n = \sum |\alpha_i|^2$ heißt Norm von $\mathfrak{a}$. Die Norm des Nullvektors ist $=0$, diejenige jedes anderen Vektors aber >0. Die nichtnegative Quadratwurzel $\sqrt{\sum|\alpha_i|^2}$ nennt man den Betrag $|\mathfrak{a}|$ von $\mathfrak{a}$. Der Vektor heißt normiert, wenn die Norm $=1$ ist. Ist $\mathfrak{a} \neq \mathfrak{o}$, d. h. $|\mathfrak{a}| > 0$, so ist der Vektor $\frac{\mathfrak{a}}{|\mathfrak{a}|}$ wegen $\frac{\mathfrak{a}}{|\mathfrak{a}|}\frac{\bar{\mathfrak{a}}}{|\mathfrak{a}|} = 1$ ein normierter Vektor. Den Übergang von $\mathfrak{a}$ zu $\frac{\mathfrak{a}}{|\mathfrak{a}|}$ werden wir als Normieren von $\mathfrak{a}$ bezeichnen.

Der Betrag des Produktes zweier Vektoren wird durch die Beträge der Faktoren gemäß der sogenannten Schwarzschen Ungleichung $|\mathfrak{a}\mathfrak{b}| \leqq |\mathfrak{a}|\,|\mathfrak{b}|$ eingeschränkt. In der gewöhnlichen Schreibweise besagt sie, daß für alle Systeme von $n+n$ Zahlen die Ungleichheit $|\sum \alpha_i\beta_i|^2 \leqq (\sum|\alpha_i|^2)(\sum|\beta_i|^2)$ gilt, die wegen $|\sum\alpha_i\beta_i| \leqq \sum|\alpha_i\beta_i|$ mit $(\sum|\alpha_i\beta_i|)^2 \leqq (\sum|\alpha_i|^2)(\sum|\beta_i|^2)$ ersichtlich gleichwertig ist. Um letztere Ungleichheit zu beweisen, beachte man nur, daß die quadratische Funktion

$$Q(X) = \sum_i (|\alpha_i|\,X + |\beta_i|)^2 = X^2 \sum_i |\alpha_i|^2 + 2\,X \sum_i |\alpha_i\beta_i| + \sum_i |\beta_i|^2$$

für reelle X durchweg $\geqq 0$ ist, so daß die Parabel $Q = Q(X)$ die X-Achse nicht durchschneiden kann und daher die Gleichung $Q(X) = 0$ entweder zwei nichtreelle Wurzeln, oder eine reelle Doppelwurzel hat. Die hierzu erforderliche Diskriminantenbedingung ist aber mit der zu beweisenden Ungleichheit offenbar identisch. — Die Schwarzsche Ungleichheit besagt, daß, wenn $\mathfrak{x}$ ein normierter Vektor ist, $|\mathfrak{a}\mathfrak{x}| \leqq |\mathfrak{a}|$ ausfällt. Mehr läßt sich nicht behaupten. Es gibt nämlich zu jedem $\mathfrak{a}$ einen normierten Vektor $\mathfrak{x}$ derart, daß $|\mathfrak{a}\mathfrak{x}| = |\mathfrak{a}|$ wird. Ist $\mathfrak{a} = \mathfrak{o}$,

so ist diese Behauptung trivial, weil dann jedes normierte $\mathfrak{x}$ von der gewünschten Beschaffenheit ist. Ist aber $\mathfrak{a} \neq \mathfrak{o}$, so kann man z. B. $\mathfrak{x} = \frac{\bar{\mathfrak{a}}}{|\mathfrak{a}|}$ setzen, um $|\mathfrak{a}\mathfrak{x}| = |\mathfrak{a}|$ zu erhalten (da ja $|\mathfrak{a}\bar{\mathfrak{a}}| = |\mathfrak{a}|^2$ ist). — Wir werden diese Bemerkung als die Umkehrung der Schwarzschen Ungleichung bezeichnen.

Ist $\{\mathfrak{a}_i\} = (\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n)$ ein geordnetes System von n Vektoren $\mathfrak{a}_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})$, so kann man $\{\mathfrak{a}_i\}$ anstatt der Matrix $\|\alpha_{ik}\|$ betrachten und umgekehrt. So ist z. B. $\{\mathfrak{e}_i\} = \|\delta_{ik}\| = \mathfrak{E}$. Ferner ist die Matrix $\|u_{ik}\|$ dann und nur dann unitär, wenn $\mathfrak{u}_i \bar{\mathfrak{u}}_k = \delta_{ik}$ gilt, so daß der Vektor $\mathfrak{u}_i = (u_{i1}, u_{i2}, \ldots, u_{in})$ normiert und auf den übrigen $n-1$ Vektoren $\mathfrak{u}_k$ orthogonal ist; $i = 1, 2, \ldots, n$. — Die m Vektoren $\mathfrak{b}_1, \mathfrak{b}_2, \ldots, \mathfrak{b}_m$, wobei $\mathfrak{b}_i = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{in})$ und zunächst $m \gtreqless n$ ist, heißen linear unabhängig, wenn eine homogen-lineare Relation $c_1 \mathfrak{b}_1 + \cdots + c_m \mathfrak{b}_m = \mathfrak{o}$ (= Nullvektor) nur mit dem trivialen Multiplikatorensystem $c_1 = \cdots = c_m = 0$ möglich ist. Im Falle $m = n$ ist für die lineare Abhängigkeit notwendig und hinreichend die Existenz von n nicht durchweg verschwindenden Zahlen c_i mit

$$c_1 \beta_{1i} + c_2 \beta_{2i} + \cdots + c_n \beta_{ni} = 0; \quad i = 1, 2, \ldots, n,$$

d.h. das Verschwinden von $\det \|\beta_{ik}\|$. Daraus folgt, daß n normierte und paarweise orthogonale Vektoren $\mathfrak{u}_i$ gewiß linear unabhängig sind. Denn die Matrix $\|u_{ik}\|$ hat als unitäre Matrix gewiß eine Reziproke, so daß $\det \|u_{ik}\| \neq 0$ wird. [Ohne die dabei benutzten besonderen Eigenschaften der *endlichen* Gleichungssysteme heranzuziehen, kann man dies nach einer klassischen Schlußweise auch wie folgt einsehen: Aus $c_1 \mathfrak{u}_1 + c_2 \mathfrak{u}_2 + \cdots + c_n \mathfrak{u}_n = \mathfrak{o}$ folgt nach Multiplikation mit $\bar{\mathfrak{u}}_i$ wegen $\bar{\mathfrak{u}}_i \mathfrak{u}_k = \delta_{ik}$ offenbar $c_i \mathfrak{u}_i \bar{\mathfrak{u}}_i = 0$, also wegen $\mathfrak{u}_i \bar{\mathfrak{u}}_i = 1$ auch $c_i = 0$; $i = 1, 2, \ldots, n$.] — Ein System $\mathfrak{a}_1, \mathfrak{a}_2, \ldots$ von linear unabhängigen Vektoren kann nicht den Nullvektor enthalten. Denn wäre etwa $\mathfrak{a}_1 = \mathfrak{o}$, so wäre $1 . \mathfrak{a}_1 + 0 . \mathfrak{a}_2 + \cdots = \mathfrak{o}$, was eine lineare Abhängigkeit ausdrückt, weil $c_1 \neq 0$ ist. — Ist $\mathfrak{a}_1 \neq \mathfrak{o}$ und von der Dimension n, so gibt es $n-1$ Vektoren $\mathfrak{a}_2, \mathfrak{a}_3, \ldots, \mathfrak{a}_n$ derart, daß die n Vektoren $\mathfrak{a}_i$ linear unabhängig sind. Denn zunächst ist $\mathfrak{a}_1 \neq \mathfrak{o}$, also hat $\mathfrak{a}_1 = (\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1n})$ mindestens eine von Null verschiedene Komponente. Es sei etwa $\alpha_{11} \neq 0$. Wir können dann z. B. $\mathfrak{a}_i = \mathfrak{e}_i$; $i = 2, 3, \ldots, n$ setzen, um ein System von n linear unabhängigen Vektoren zu erhalten. Für $i > 1$ ist dann nämlich $\alpha_{ik} = \delta_{ik}$, so daß $\det \|\alpha_{ik}\|$, wie man sich bei Entwicklung nach der letzten Zeile durch vollständige Induktion leicht überzeugt, $= \alpha_{11}$, also $\neq 0$ sein wird. — Hingegen sind $n+1$ Vektoren $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n, \mathfrak{a}_{n+1}$ der Dimension n immer linear abhängig. Bezeichnet nämlich $\mathfrak{c}_i$ den

$(n+1)$-dimensionalen Vektor, dessen n ersten Komponenten mit denjenigen von $\mathfrak{a}_i$ identisch sind, während $c_{i\,n+1}=0$ ist, so gilt $\det\|c_{ik}\|=0$, da alle Elemente der letzten Kolonne verschwinden. Mithin sind die $n+1$ Vektoren $\mathfrak{c}_i$ der Dimension $n+1$ linear abhängig. Erst recht besteht also eine homogen-lineare Beziehung zwischen den $n+1$ Vektoren $\mathfrak{a}_i$ der Dimension n, w. z. b. w.

§ 6. Der Schmidtsche Orthogonalisierungsprozeß.

Wir haben vorhin gesehen, daß n paarweise orthogonale Vektoren, sobald sie normiert sind oder wenigstens normiert werden können (d. h. wenn keiner dieser Vektoren der Nullvektor ist), gewiß linear unabhängig sind. Nach E. Schmidt kann man diesen Satz gewissermaßen umkehren; dazu wollen wir nun übergehen. Eine Matrix $\mathfrak{L}=\|l_{ik}\|$ möge kanonisch genannt werden, wenn darin für $i<k$ $l_{ik}=0$ gilt, so daß

$$\mathfrak{L}=\left\|\begin{matrix} l_{11} & 0 & 0 & \ldots & 0 \\ l_{21} & l_{22} & 0 & \ldots & 0 \\ l_{31} & l_{32} & l_{33} & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \ldots & l_{nn} \end{matrix}\right\| \tag{14}$$

ist (wobei auch die in oder unterhalb der Diagonale gelegenen Elemente verschwinden *dürfen*). Z. B. die Diagonalmatrizen sind gewiß kanonisch. Man ersieht durch Entwicklung von $\det\|l_{ik}\|$ nach der ersten Zeile, daß $\det\|l_{ik}\|=l_{11}l_{22}\cdots l_{nn}$, also insbesondere dann und nur dann $\neq 0$ ist, wenn kein l_{ii} verschwindet. Ist diese Bedingung erfüllt, so wollen wir $\|l_{ik}\|$ eine rekursive Matrix nennen. Die rekursiven Matrizen sind also die nicht entarteten kanonischen Matrizen. Ist $\|r_{ik}\|$ kanonisch und setzt man

$$\mathfrak{q}_i=\sum_{k=1}^{n} r_{ik}\,\mathfrak{a}_k; \qquad i=1,2,\ldots,n, \tag{15}$$

so hängt $\mathfrak{q}_j$ nur von den j ersten Vektoren $\mathfrak{a}_i$ ab. Ist überdies $\|r_{ik}\|$ rekursiv, so kann man umgekehrt $\mathfrak{a}_j$ durch die j ersten Vektoren $\mathfrak{q}_i$, $i=1,2,\ldots,j$ $(\leqq n)$ homogen-linear ausdrücken, indem aus dem System (15), wegen $r_{ik}=0$, $i<k$ und wegen $r_{ii}\neq 0$, die Vektoren $\mathfrak{a}_1,\mathfrak{a}_2,\ldots$ *rekursiv* berechnet werden können. — Der Schmidtsche Satz besagt nun, daß zu jedem System $\mathfrak{a}_1,\mathfrak{a}_2,\ldots,\mathfrak{a}_n$ linear-unabhängiger Vektoren der Dimension n eine — übrigens nicht nur eine — rekursive Matrix $\|r_{ik}\|$ angegeben werden kann derart, daß das durch (15) er-

klärte Vektorensystem normiert und orthogonal ist, $\mathfrak{q}_i \bar{\mathfrak{q}}_k = \delta_{ik}$. Und zwar kann dabei $\mathfrak{q}_1 = \mathfrak{a}_1$ angenommen werden, sobald die hierfür notwendige Bedingung $|\mathfrak{a}_1| = 1$ erfüllt ist.

Zunächst beweisen wir folgenden Hilfssatz: Sind $\mathfrak{a}$ und $\mathfrak{b}$ zwei linear unabhängige Vektoren, so gibt es eine Zahl γ derart, daß die beiden Vektoren $\mathfrak{a}$, $\mathfrak{b} + \gamma\mathfrak{a}$ orthogonal und linear unabhängig sind. Und zwar ist $\gamma = -\frac{\bar{\mathfrak{a}}\mathfrak{b}}{|\mathfrak{a}|^2}$ zu setzen. Der Nenner ist dabei von Null verschieden, da das System $\mathfrak{a}, \mathfrak{b}$ von linear unabhängigen Vektoren den Nullvektor nicht enthalten kann. — Man hat mit dem obigen Wert von γ offenbar $\bar{\mathfrak{a}}(\mathfrak{b} + \gamma\mathfrak{a}) = \bar{\mathfrak{a}}\mathfrak{b} + \gamma|\mathfrak{a}|^2 = 0$, so daß $\mathfrak{a}$ und $\mathfrak{b} + \gamma\mathfrak{a}$ orthogonal sind. Sie sind ferner linear unabhängig. Aus $c_1\mathfrak{a} + c_2\mathfrak{b} + c_2\gamma\mathfrak{a} = \mathfrak{o}$ folgt nämlich in Anbetracht der linearen Unabhängigkeit von $\mathfrak{a}$ und $\mathfrak{b}$ zunächst $c_1 + c_2\gamma = 0$, $c_2 = 0$ und daraus $c_1 = 0$, $c_2 = 0$. Damit ist der Hilfssatz bewiesen.

Nehmen wir nun an, daß wir aus den n linear unabhängigen Vektoren $\mathfrak{a}_i$ bereits $p\,(< n)$ Vektoren $\mathfrak{r}_1, \mathfrak{r}_2, \ldots, \mathfrak{r}_p$ den folgenden Bedingungen gemäß abgeleitet haben: 1) sie sind orthogonal, $\mathfrak{r}_i \bar{\mathfrak{r}}_k = 0$, $i \neq k$; 2) $\mathfrak{r}_j$ ist eine homogene lineare Verbindung der Vektoren $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_j$, in welcher der Koeffizient von $\mathfrak{a}_j$ gleich 1 ist; $j = 1, 2, \ldots, p$. Für $p = 2$ kann dies nach dem Hilfssatz gewiß erreicht werden, und wir wollen von p auf $p + 1$ schließen. Zu diesem Ende bemerken wir zunächst, daß $\mathfrak{a}_{p+1}$ von $\mathfrak{r}_i$ gewiß linear unabhängig ist; $i = 1, 2, \ldots, p$. Aus $c_1\mathfrak{a}_{p+1} + c_2\mathfrak{r}_i = \mathfrak{o}$ folgt nämlich, wegen $i < p + 1$ und da $\mathfrak{r}_i$ nur von den i ersten Vektoren $\mathfrak{a}_k$ abhängt, offenbar $c_1 = 0$, da sonst $\mathfrak{a}_{p+1} + \frac{c_2}{c_1}\mathfrak{r}_i = \mathfrak{o}$ eine lineare Abhängigkeit zwischen den $i + 1$ Vektoren $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_i; \mathfrak{a}_{p+1}$ ausdrücken würde [vgl. 2)], während nach Voraussetzung die n Vektoren $\mathfrak{a}$ linear unabhängig sind. Es ist also $c_2\mathfrak{r}_i = \mathfrak{o}$. Wir haben zu beweisen, daß außer c_1 auch $c_2 = 0$ ist. Wäre dies nicht der Fall, so müßte $\mathfrak{r}_i = \mathfrak{o}$ sein. Aus $\mathfrak{r}_i = \mathfrak{o}$ folgt aber wieder mit Rücksicht auf die Annahme 2) eine lineare Abhängigkeit zwischen den $\mathfrak{a}$. — Damit ist gezeigt, daß $\mathfrak{a}_{p+1}$ von $\mathfrak{r}_i$ linear unabhängig ist. Nach dem Hilfssatz gibt es also eine Zahl $\gamma_{j\,p+1}$ derart, daß die beiden Vektoren $\mathfrak{r}_j$, $\mathfrak{a}_{p+1} + \gamma_{j\,p+1}\mathfrak{r}_j$ orthogonal und linear unabhängig sind. Setzen wir nun $\mathfrak{r}_{p+1} = \mathfrak{a}_{p+1} + \sum_{i=1}^{p} \gamma_{i\,p+1}\mathfrak{r}_i$, so ist $\mathfrak{r}_{p+1}$ eine homogen-lineare Verbindung der $p + 1$ ersten $\mathfrak{a}_i$ und der Koeffizient von $\mathfrak{a}_{p+1}$ ist gleich 1, und zwar wegen 2). Es gilt ferner für $j \leqq p$ wegen 1)

$$\bar{\mathfrak{r}}_j\mathfrak{r}_{p+1} = \bar{\mathfrak{r}}_j\mathfrak{a}_{p+1} + \sum_{i=1}^{p} \gamma_{i\,p+1}\bar{\mathfrak{r}}_j\mathfrak{r}_i = \bar{\mathfrak{r}}_j\mathfrak{a}_{p+1} + \gamma_{j\,p+1}\bar{\mathfrak{r}}_j\mathfrak{r}_j.$$

Also ist $\bar{\mathfrak{r}}_j \mathfrak{r}_{p+1} = 0$, da $\bar{\mathfrak{r}}_j(\mathfrak{a}_{p+1} + \gamma_{j\,p+1}\mathfrak{r}_j)$ nach der Definition von $\gamma_{j\,p+1}$ verschwindet. Damit ist die vollständige Induktion beendet.

Wir haben damit eine rekursive Matrix $\| l_{ik} \|$ derart, daß die n Vektoren $\mathfrak{r}_i = \sum_{k=1}^{n} l_{ik}\,\mathfrak{a}_k$ orthogonal sind, und es ist dabei $l_{ii} = 1$, also insbesondere $\mathfrak{r}_1 = \mathfrak{a}_1$. Es ist $\mathfrak{r}_i \neq \mathfrak{o}$, da die gegenteilige Annahme, wie wir gesehen haben, zu einem Widerspruch zu 2) führt. Wir können daher die Vektoren $\mathfrak{r}_i$ durch Division mit $|\mathfrak{r}_i|$ normieren, wodurch weder die Orthogonalität noch der rekursive Charakter verloren geht. — Damit ist der Schmidtsche Satz bewiesen, wir können ja in (15) einfach $\mathfrak{q}_i = \frac{\mathfrak{r}_i}{|\mathfrak{r}_i|}$, $r_{ik} = \frac{l_{ik}}{|\mathfrak{r}_i|}$ setzen.

§ 7. Die Matrizentransformation.

Es sei $\mathfrak{A} = \| a_{ik} \|$ eine Matrix und $\mathfrak{x} = (x_1, \ldots, x_n)$ ein Vektor. Wir wollen sagen, der Vektor $\mathfrak{c}$, dessen Komponenten $c_i = \sum_{k=1}^{n} a_{ik} x_k$ sind, entstehe aus $\mathfrak{x}$ durch die Transformation $\mathfrak{A}$; kurz können wir $\mathfrak{c} = \mathfrak{A}\mathfrak{x}$ schreiben (hingegen ist das „Produkt" $\mathfrak{x}\mathfrak{A}$ nicht erklärt). Dementsprechend wollen wir die Transformation einer Matrix $\mathfrak{A}$ durch eine transformierende Matrix $\mathfrak{S}$ erklären. Wir werden dabei voraussetzen, daß $\mathfrak{S}$ nicht entartet ist, während dies für $\mathfrak{A}$ nicht zuzutreffen braucht. — Wir transformieren die beiden Vektoren $\mathfrak{x}$, $\mathfrak{c} = \mathfrak{A}\mathfrak{x}$ durch die Matrix $\mathfrak{S}$ in $\mathfrak{S}\mathfrak{x}$, $\mathfrak{S}\mathfrak{c}$ und fragen nach einer Matrix $\mathfrak{T}$, durch welche $\mathfrak{S}\mathfrak{x}$ in $\mathfrak{S}\mathfrak{c}$ transformiert wird. Wir werden $\mathfrak{T}$ sinngemäß als eine Matrix bezeichnen können, in welche $\mathfrak{A}$ durch $\mathfrak{S}$ transformiert wird. — Werden auf den Vektor $\mathfrak{S}\mathfrak{x}$ nacheinander die Transformationen $\mathfrak{S}^{-1}$, $\mathfrak{A}$, $\mathfrak{S}$ ausgeübt (in dieser Reihenfolge!), so ergeben sich nacheinander die Vektoren

$$\mathfrak{S}^{-1}(\mathfrak{S}\mathfrak{x}) = \mathfrak{x}, \quad \mathfrak{A}(\mathfrak{S}^{-1}(\mathfrak{S}\mathfrak{x})) = \mathfrak{A}\mathfrak{x} = \mathfrak{c}, \quad \mathfrak{S}(\mathfrak{A}(\mathfrak{S}^{-1}(\mathfrak{S}\mathfrak{x}))) = \mathfrak{S}\mathfrak{c}.$$

Nun ist aber, wie man leicht verifiziert, $\mathfrak{C}(\mathfrak{D}\mathfrak{a}) = (\mathfrak{C}\mathfrak{D})\mathfrak{a}$, unter $(\mathfrak{C}\mathfrak{D})$ das Produkt zweier Matrizen, unter $(\mathfrak{C}\mathfrak{D})\mathfrak{a}$ die Transformation von $\mathfrak{a}$ durch $(\mathfrak{C}\mathfrak{D})$, endlich unter $\mathfrak{C}(\mathfrak{D}\mathfrak{a})$ die Transformation des Vektors $\mathfrak{D}\mathfrak{a}$ durch die Matrix $\mathfrak{C}$ verstanden. Man hat daher $\mathfrak{S}\mathfrak{c} = (\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})(\mathfrak{S}\mathfrak{x})$, was freilich, wegen $\mathfrak{c} = \mathfrak{A}\mathfrak{x}$ und da $\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}\mathfrak{S} = \mathfrak{S}\mathfrak{A}$ ist, jetzt auch direkt einleuchtet. Wir wollen festsetzen: eine Matrix $\mathfrak{A}$ mit einer Matrix $\mathfrak{S}$ zu transformieren, bedeutet das Ersetzen von $\mathfrak{A}$ durch $\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$, wobei also $\det\mathfrak{S} \neq 0$ angenommen wird, nicht aber auch $\det\mathfrak{A} \neq 0$. — Wir können den Inhalt dieser Betrachtungen auch wie folgt deuten. Ist $\mathfrak{x} = (x_1, \ldots, x_n)$ eine Lösung des (homogenen oder

inhomogenen) Gleichungssystems $\sum a_{ik} x_k = c_i$, das in unserer jetzigen Schreibweise als $\mathfrak{A}\mathfrak{x} = \mathfrak{c}$ geschrieben werden kann, so ist der Vektor $\mathfrak{y} = \mathfrak{S}\mathfrak{x}$ gewiß eine Lösung des Gleichungssystems $(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})\mathfrak{y} = \mathfrak{S}\mathfrak{c}$. — Wegen $(\mathfrak{S}^{-1})^{-1} = \mathfrak{S}$ und $\mathfrak{S}\mathfrak{S}^{-1} = \mathfrak{E}$ ist $\mathfrak{S}^{-1}(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})(\mathfrak{S}^{-1})^{-1} = \mathfrak{A}$, d. h. $\mathfrak{A}$ entsteht aus $\mathfrak{B}$ durch die Transformation $\mathfrak{S}^{-1}$, wenn $\mathfrak{B}$ aus $\mathfrak{A}$ durch die Transformation $\mathfrak{S}$ entsteht. — Ist $\mathfrak{A}^{-1}$ vorhanden, so ist wegen $(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})(\mathfrak{S}\mathfrak{A}^{-1}\mathfrak{S}^{-1}) = \mathfrak{E}$ auch $(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})^{-1}$ vorhanden, und zwar gleich $\mathfrak{S}\mathfrak{A}^{-1}\mathfrak{S}^{-1}$, so daß die Matrix $\mathfrak{A}$ und ihre Reziproke sich *kogredient*, d. h. durch eine und dieselbe Transformation transformieren. Offenbar können die Determinanten von $\mathfrak{A}$ und $\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$ nur zugleich verschwinden. Genauer gilt für alle Fälle $\det\mathfrak{A} = \det\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$, die Determinante ist Matrizentransformationen gegenüber invariant. Man hat nämlich nach dem Multiplikationssatz der Determinanten $\det(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}) = (\det\mathfrak{S})(\det\mathfrak{A})(\det\mathfrak{S}^{-1})$, und es ist, wie erwähnt, $(\det\mathfrak{S})(\det\mathfrak{S}^{-1}) = 1$. — Sind die Vektoren $\mathfrak{x}_1, \mathfrak{x}_2, \ldots$ linear unabhängig, so sind es auch die transformierten Vektoren $\mathfrak{S}\mathfrak{x}_1, \mathfrak{S}\mathfrak{x}_2, \ldots$, (und freilich auch umgekehrt, da auch $\mathfrak{S}^{-1}$ ein $\mathfrak{S}$ ist). Genauer, aus $\sum c_i \mathfrak{x}_i = \mathfrak{o}$ folgt nach Transformation mit $\mathfrak{S}$ wegen $\mathfrak{S}\mathfrak{o} = \mathfrak{o}$ offenbar $\sum c_i \mathfrak{S}\mathfrak{x}_i = \mathfrak{o}$. Insbesondere haben daher die beiden homogenen Gleichungssysteme $\mathfrak{A}\mathfrak{x} = \mathfrak{o}$, $(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})\mathfrak{x} = \mathfrak{o}$ dieselbe Anzahl $(\geqq 0)$ linear unabhängiger, von $\mathfrak{o}$ verschiedener Lösungen.

Es sei noch erwähnt, daß $F(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}) = \mathfrak{S}F(\mathfrak{A})\mathfrak{S}^{-1}$ gilt, wenn $F(X)$ etwa ein Polynom bezeichnet [im Falle $F(X) = X^{-1}$ haben wir die Richtigkeit dieser Beziehung bereits früher vermerkt]. Wegen $\mathfrak{S}(c_1\mathfrak{A}_1 + c_2\mathfrak{A}_2)\mathfrak{S}^{-1} = c_1\mathfrak{S}\mathfrak{A}_1\mathfrak{S}^{-1} + c_2\mathfrak{S}\mathfrak{A}_2\mathfrak{S}^{-1}$ können wir uns dabei auf den Fall $F(X) = X^m$, $m = 0, 1, 2, \ldots$ beschränken. Für $m = 0$ und für $m = 1$ ist dann die Behauptung trivial. Es gilt aber wegen $\mathfrak{S}^{-1}\mathfrak{S} = \mathfrak{E}$ auch

$$\mathfrak{S}\mathfrak{A}^2\mathfrak{S}^{-1} = \mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1} = (\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})^2,$$

$$\mathfrak{S}\mathfrak{A}^3\mathfrak{S}^{-1} = \mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1} = (\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})^3,$$

usf.

§ 8. Das Spektrum.

Bei jedem Wert des Parameters λ ist $\mathfrak{S}(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{S}^{-1} = (\lambda\mathfrak{S} - \mathfrak{S}\mathfrak{A})\mathfrak{S}^{-1} = \lambda\mathfrak{E} - \mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$, woraus wegen $\det(\mathfrak{S}\mathfrak{B}\mathfrak{S}^{-1}) = \det\mathfrak{B}$ offenbar $\det(\lambda\mathfrak{E} - \mathfrak{A}) = \det(\lambda\mathfrak{E} - \mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})$, d. h. die Invarianz von $\det(\lambda\mathfrak{E} - \mathfrak{A})$ gegenüber Matrizentransformationen von $\mathfrak{A}$ folgt. Dies beruht darauf, daß sich die beiden Matrizen $\mathfrak{A}$, $\lambda\mathfrak{E} - \mathfrak{A}$ kogredient transformieren. Man nennt $\det(\lambda\mathfrak{E} - \mathfrak{A})$ charakteristisches oder Säkularpolynom von $\mathfrak{A}$. Es ist in λ eine rationale ganze Funktion

vom Grade n, worin der Koeffizient von λ^n gleich 1 ist. Unter den $n!$ Termen von $\det\|\lambda\,\delta_{ik} - a_{ik}\|$ gibt es nämlich einen und nur einen derart, daß λ in jedem seiner n Faktoren enthalten ist, und dieser Term ist das Produkt $\prod(\lambda - a_{ii}) = \lambda^n - \lambda^{n-1}\sum a_{ii} + \cdots$ der diagonalen Elemente. Die n voneinander nicht notwendig verschiedenen Wurzeln $\lambda_1, \lambda_2, \ldots, \lambda_n$ der Säkulargleichung $\det(\lambda\,\mathfrak{E} - \mathfrak{A}) = 0$ heißen Eigenwerte von $\mathfrak{A}$, ihre Gesamtheit nennt man nach einem ebenfalls physikalischen, auf Wirtinger zurückgehenden und seit Hilbert (1906) eingebürgerten Sprachgebrauch das Spektrum von $\mathfrak{A}$. Das Spektrum ist Matrizentransformationen gegenüber invariant, da bereits das Säkularpolynom invariant ist, also erst recht die Säkulargleichung. Da die beiden Gleichungen $\det(0\,\mathfrak{E} - \mathfrak{A}) = 0$, $\det\mathfrak{A} = 0$ gleichwertig sind, so ist das Verschwinden mindestens eines Eigenwertes notwendig und hinreichend dafür, daß $\mathfrak{A}$ entartet ist. Ist dies nicht der Fall, so folgt aus $\det(\lambda\,\mathfrak{E} - \mathfrak{A}) = 0$ wegen $\det(\lambda\,\mathfrak{E} - \mathfrak{A}) = [\det\mathfrak{A}]\,[\det(\lambda\,\mathfrak{A}^{-1} - \mathfrak{E})]$ offenbar $\det(\lambda^{-1}\,\mathfrak{E} - \mathfrak{A}^{-1}) = 0$ und umgekehrt, so daß die reziproken Werte der Eigenwerte von $\mathfrak{A}$ die Eigenwerte von $\mathfrak{A}^{-1}$ sind.

Die Eigenwerte von $\mathfrak{A}$ sind diejenigen Zahlen λ, bei welchen die zu $\lambda\,\mathfrak{E} - \mathfrak{A}$ gehörigen homogenen Gleichungen eine Nullösung zulassen, die diesmal als eine zu λ gehörige Eigenlösung genannt wird. Die Anzahl der zu λ gehörigen linear unabhängigen Lösungen ist, nach dem Vorhergehenden, Matrizentransformationen gegenüber invariant.

Wir wollen jetzt (im wesentlichen nach Borchardt) die Eigenwerte und die Eigenlösungen der Funktionen von $\mathfrak{A}$ untersuchen. Betrachten wir zunächst die positiven ganzen Potenzen von $\mathfrak{A}$ und nehmen wir einstweilen an, daß λ ein von Null verschiedener Eigenwert von $\mathfrak{A}$ und $\mathfrak{x}$ eine zugehörige Eigenlösung ist. Wegen $(\lambda\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{x} = \mathfrak{o}$ oder $(\lambda\,\mathfrak{E})\,\mathfrak{x} = \mathfrak{A}\,\mathfrak{x}$ gilt dann $\mathfrak{x} = \lambda^{-1}\,\mathfrak{A}\,\mathfrak{x}$, also auch $\mathfrak{x} = \lambda^{-1}\,\mathfrak{A}(\lambda^{-1}\,\mathfrak{A}\,\mathfrak{x}) = \lambda^{-2}\,\mathfrak{A}^2\,\mathfrak{x}$, und allgemeiner $\mathfrak{x} = \lambda^{-n}\,\mathfrak{A}^n\,\mathfrak{x}$, d. h. $(\lambda^n\,\mathfrak{E} - \mathfrak{A}^n)\,\mathfrak{x} = \mathfrak{o}$. — Dies gilt, auch wenn der Eigenwert $\lambda = 0$ ist. Wir haben nur zu zeigen, daß aus $\mathfrak{A}\,\mathfrak{x} = \mathfrak{o}$ folgt $\mathfrak{A}^n\,\mathfrak{x} = \mathfrak{o}$, und dies ist selbstverständlich, man braucht ja nur den Vektor $\mathfrak{A}\,\mathfrak{x}$ vermittels $\mathfrak{A}^{n-1}$ zu transformieren und dabei $\mathfrak{A}^{n-1}\,\mathfrak{o} = \mathfrak{o}$ zu beachten. — Für $n = 0$ ist $\lambda^n\,\mathfrak{E} - \mathfrak{A}^n = \mathfrak{E} - \mathfrak{E} = \mathfrak{O}$, also $(\lambda^n\,\mathfrak{E} - \mathfrak{A}^n)\,\mathfrak{x} = \mathfrak{o}$ eine Trivialität. Endlich gilt dies, sobald $\mathfrak{A}^{-1}$ überhaupt vorhanden ist, auch für $n = -1$ (also für $n = 0, \pm 1, \pm 2, \ldots$). Wir können nämlich dabei $\lambda \neq 0$ annehmen und dann folgt die Behauptung $(\lambda^{-1}\,\mathfrak{E} - \mathfrak{A}^{-1})\,\mathfrak{x} = \mathfrak{o}$ aus der Voraussetzung $(\lambda\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{x} = \mathfrak{o}$ einfach durch Transformation vermittels $\mathfrak{A}^{-1}$. Ist endlich $F(X)$ eine z. B. rationale und der Einfachheit halber zugleich ganze Funktion von X, $F(X) = C_0 + C_1\,X + C_2\,X^2 + \cdots$, so gilt, da $C_k\,(\lambda^k\,\mathfrak{E} - \mathfrak{A}^k)\,\mathfrak{x} = \mathfrak{o}$ ist, $C_k\,\lambda^k\,\mathfrak{x} = C_k\,\mathfrak{A}^k\,\mathfrak{x}$, woraus nach Summation in bezug auf k sofort

$F(\lambda)\mathfrak{x} = F(\mathfrak{A})\mathfrak{x}$ und daher auch $(F(\lambda)\mathfrak{E} - F(\mathfrak{A}))\mathfrak{x} = \mathfrak{o}$ folgt. Ist also λ ein Eigenwert von $\mathfrak{A}$ und $\mathfrak{x}$ eine zugehörige Eigenlösung, so ist $F(\lambda)$ gewiß ein Eigenwert von $F(\mathfrak{A})$ und $\mathfrak{x}$ eine zugehörige Eigenlösung. Auf die umgekehrte Frage, inwiefern Eigenwerte und Eigenlösungen von $\mathfrak{A}$ aus denjenigen von $F(\mathfrak{A})$ durch algebraische Diskussionen erhalten werden können, kann hier nicht näher eingegangen werden.

§ 9. Die Resolvente und die Hilbertsche Funktionalgleichung.

Die inhomogenen Gleichungen $\lambda x_i - \sum a_{ik} x_k = c_i$ oder $(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{c}$ sind bei jedem $\mathfrak{c}$ dann und nur dann lösbar, wenn $\det(\lambda\mathfrak{E} - \mathfrak{A}) \neq 0$, d. h. λ kein Eigenwert ist, was wir nunmehr voraussetzen. Dann hat $(\lambda\mathfrak{E} - \mathfrak{A})$ eine Reziproke $(\lambda\mathfrak{E} - \mathfrak{A})^{-1}$, die, als Funktion von λ betrachtet, auch durch $\mathfrak{R}_\lambda = \| R_{ik}(\lambda) \|$ bezeichnet und Resolvente von $\mathfrak{A}$ genannt werden soll.

Ist $\mathfrak{C}, \mathfrak{D}$ ein Paar nicht entarteter Matrizen, die vertauschbar sind, so ist $\mathfrak{C}^{-1}$ mit $\mathfrak{D}^{-1}$ vertauschbar. Für alle Fälle gilt nämlich $\mathfrak{C}\mathfrak{D}\mathfrak{D}^{-1}\mathfrak{C}^{-1} = \mathfrak{E}$, also $(\mathfrak{C}\mathfrak{D})^{-1} = \mathfrak{D}^{-1}\mathfrak{C}^{-1}$ und ebenso $(\mathfrak{D}\mathfrak{C})^{-1} = \mathfrak{C}^{-1}\mathfrak{D}^{-1}$. Ist aber $\mathfrak{C}\mathfrak{D} = \mathfrak{D}\mathfrak{C}$, also, da es nur eine Reziproke geben kann[1]), auch $(\mathfrak{C}\mathfrak{D})^{-1} = (\mathfrak{D}\mathfrak{C})^{-1}$, so folgt daraus die Behauptung, $\mathfrak{D}^{-1}\mathfrak{C}^{-1} = \mathfrak{C}^{-1}\mathfrak{D}^{-1}$. — Sind nun λ', λ'' zwei λ-Werte, so sind die beiden Matrizen $(\lambda'\mathfrak{E} - \mathfrak{A})$, $(\lambda''\mathfrak{E} - \mathfrak{A})$ gewiß vertauschbar, da ihr Produkt $\lambda'\lambda''\mathfrak{E} - (\lambda' + \lambda'')\mathfrak{A} + \mathfrak{A}^2$ in den beiden Faktoren symmetrisch ist. Folglich gilt auch $\mathfrak{R}_{\lambda'}\mathfrak{R}_{\lambda''} = \mathfrak{R}_{\lambda''}\mathfrak{R}_{\lambda'}$. — Nach Definition ist $(\lambda'\mathfrak{E} - \mathfrak{A})\mathfrak{R}_{\lambda'} = \mathfrak{E}$, also $\mathfrak{A}\mathfrak{R}_{\lambda'} = \lambda'\mathfrak{R}_{\lambda'} - \mathfrak{E}$, mithin

$$\mathfrak{A}\mathfrak{R}_{\lambda'}\mathfrak{R}_{\lambda''} = \lambda'\mathfrak{R}_{\lambda'}\mathfrak{R}_{\lambda''} - \mathfrak{R}_{\lambda''}.$$

Vertauscht man λ' mit λ'', so folgt wegen $\mathfrak{R}_{\lambda'}\mathfrak{R}_{\lambda''} = \mathfrak{R}_{\lambda''}\mathfrak{R}_{\lambda'}$ auch

$$\mathfrak{A}\mathfrak{R}_{\lambda'}\mathfrak{R}_{\lambda''} = \lambda''\mathfrak{R}_{\lambda'}\mathfrak{R}_{\lambda''} - \mathfrak{R}_{\lambda'}.$$

Der Vergleich dieser beiden Formeln ergibt die Hilbertsche Funktionalgleichung

$$(\lambda' - \lambda'')\mathfrak{R}_{\lambda'}\mathfrak{R}_{\lambda''} = \mathfrak{R}_{\lambda''} - \mathfrak{R}_{\lambda'}.$$

Wendet man die Regel (9) anstatt von $\mathfrak{A}$ auf die Matrix $\lambda\mathfrak{E} - \mathfrak{A}$ an, so folgt

(16) $$R_{ik}(\lambda) = \frac{S_{ik}(\lambda)}{D(\lambda)},$$

[1]) Daß es eine Reziproke gibt, folgt bereits daraus, daß wegen $\det\mathfrak{C} \neq 0$ und $\det\mathfrak{D} \neq 0$ auch $\det(\mathfrak{C}\mathfrak{D}) = (\det\mathfrak{C})(\det\mathfrak{D}) \neq 0$ ist.

wobei $S_{ik}(\lambda)$ eine rationale ganze Funktion von einem Grade $\leqq n-1$ und $D(\lambda)$ das Säkularpolynom, also eine rationale ganze Funktion vom Grade n bezeichnet. Sind daher $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(p)}$ die *verschiedenen* Eigenwerte von $\mathfrak{A}$, so daß, wenn $\lambda^{(j)}$ eine s_j-fache Wurzel von $D(\lambda) = 0$ ist,

$$(17) \qquad s_1 + s_2 + \cdots + s_p = n \qquad (s_j \geqq 1,\ 1 \leqq p \leqq n)$$

wird, so gilt eine Partialbruchzerlegung

$$(18) \qquad R_{ik}(\lambda) = \sum_{j=1}^{p} \sum_{h=1}^{s_j} \frac{\varrho_{ik}^{hj}}{(\lambda - \lambda^{(j)})^h}; \qquad i, k = 1, 2, \ldots, n,$$

wobei die Zahlen ϱ_{ik}^{hj} eindeutig bestimmt sind. Es verhält sich also jede Funktion $R_{ik}(\lambda)$ an der Stelle λ gewiß regulär, wenn λ kein Eigenwert ist (und zwar auch für $\lambda = \infty$). Es kann jedoch ein $R_{ik}(\lambda)$ auch an einer Stelle $\lambda^{(j)}$ regulär sein, d. h. es können bei festem j, i und k alle s_j Zahlen $\varrho_{ik}^{1j}, \varrho_{ik}^{2j}, \ldots, \varrho_{ik}^{s_j j}$ verschwinden. Ist z. B. $\mathfrak{A} = \|a_{ik}\| = \|\delta_{ik}\lambda_i\|$ und sind die n Zahlen λ_i voneinander verschieden, so ist, wie man leicht verifiziert (vgl. § 13), $R_{ik}(\lambda) = \frac{\delta_{ik}}{\lambda - \lambda_i}$, so daß die λ_i Eigenwerte sind; an der Stelle $\lambda = \lambda_i$ ist nur $R_{ii}(\lambda)$ singulär, während die übrigen $n^2 - 1$ Funktionen $R_{ik}(\lambda)$ sich daselbst regulär verhalten. Für alle Fälle ist aber, wie wir jetzt zeigen wollen, für $\lambda = \lambda^*$ *mindestens ein* $R_{ik}(\lambda)$ singulär, wenn λ^* ein Eigenwert ist. [Da an den Stellen, die keinem Eigenwerte entsprechen, jedes $R_{ik}(\lambda)$ gewiß regulär ist, so können wir dies auch so ausdrücken: das Spektrum von $\mathfrak{A}$ ist mit der Gesamtheit der *singulären Stellen der Matrix* $(\lambda\mathfrak{E} - \mathfrak{A})^{-1} = \|R_{ik}(\lambda)\|$ identisch. Und zwar sind alle Singularitäten von rationalem Charakter. Bei unendlichen Matrizen wird dies nicht mehr zutreffen.]

Nehmen wir zwecks Zurückführung ad absurdum an, daß an der Stelle λ_* sich alle n^2 Funktionen $R_{ik}(\lambda)$ regulär verhalten, obwohl λ_* ein Eigenwert ist. Da es nur endlich viele Eigenwerte gibt, können wir eine Zahlenfolge $\lambda^{[1]}, \lambda^{[2]}, \ldots, \lambda^{[m]}, \ldots$ derart bestimmen, daß $\lambda^{[m]} \to \lambda_*$, $m \to +\infty$ gilt, und kein $\lambda^{[m]}$ ein Eigenwert ist, so daß $\|R_{ik}(\lambda^{[m]})\|$ reziprok zu $\|\lambda^{[m]}\delta_{ik} - a_{ik}\|$, also das Produkt dieser beiden Matrizen $= \|\delta_{ik}\|$, d. h.

$$(19) \qquad \sum_{\nu=1}^{n} R_{i\nu}(\lambda^{[m]}) \cdot (\lambda^{[m]}\delta_{\nu k} - a_{\nu k}) = \delta_{ik}; \qquad i, k = 1, 2, \ldots, n$$

wird. Da nach Voraussetzung für $\lambda = \lambda_*$ alle n^2 Funktionen $R_{ik}(\lambda)$ regulär sind, so strebt jedes $R_{ik}(\lambda^{[m]})$ für $\lambda^{[m]} \to \lambda_*$ der Grenze

$R_{ik}(\lambda_*)$ zu. Die Vornahme des Grenzüberganges ergibt aber aus (19), daß

$$\sum_{\nu=1}^{n} R_{i\nu}(\lambda_*) \cdot (\lambda_* \delta_{\nu k} - a_{\nu k}) = \delta_{ik}; \qquad i, k = 1, 2, \ldots, n$$

gilt, und daher $\| \lambda_* \delta_{ik} - a_{ik} \|$ eine Reziproke hat, nämlich die Matrix $\| R_{ik}(\lambda_*) \|$. Mithin müßte $\det \| \lambda_* \delta_{ik} - a_{ik} \| \neq 0$ sein, d. h. λ_* wäre keine Wurzel der Säkulargleichung, im Widerspruch zu der Voraussetzung.

§ 10. Die C. Neumannschen Reihen.

Da im Punkte $\lambda = \infty$ wegen (18) alle $R_{ik}(\lambda)$ regulär verschwinden, so gilt dort eine Taylorsche Entwicklung

$$(20) \qquad R_{ik}(\lambda) = \sum_{m=0}^{\infty} \frac{a_{ik}^{(m)}}{\lambda^{m+1}}; \qquad i, k = 1, 2, \ldots, n,$$

und zwar sind *alle* n^2 Reihen für $|\lambda| > \varrho$ und *nur* für $|\lambda| > \varrho$ konvergent, wenn ϱ den Maximalbetrag der Eigenwerte von $\mathfrak{A}$ bedeutet. Sind alle Eigenwerte $= 0$, so ist auch $\varrho = 0$ und die Reihen sind konvergent, sobald $\lambda \neq 0$ ist. Es ist dabei entweder mindestens ein $a_{ik}^{(m)} \neq 0$ und dann ist die Reihe von $R_{ik}(\lambda)$ für $\lambda = 0$ sinnlos, oder es sind alle $a_{ik}^{(m)} = 0$. Letzter Fall ist aber ausgeschlossen. Sonst wäre nämlich $R_{ik}(\lambda) \equiv 0$ für alle i und k, d. h. $\mathfrak{R}_\lambda \equiv \mathfrak{O}$, und daher $\mathfrak{R}_\lambda(\lambda \mathfrak{E} - \mathfrak{A}) \equiv \mathfrak{O}$, während das Produkt für hinreichend große $|\lambda|$ gewiß $\equiv \mathfrak{E}$ sein muß.

Setzen wir $\| a_{ik}^{(m)} \| = \mathfrak{A}^{(m)}$, so können wir für (20) auch $\mathfrak{R}_\lambda = \sum_{m=0}^{\infty} \frac{\mathfrak{A}^{(m)}}{\lambda^{m+1}}$ setzen, wobei die Matrizenpotenzreihe (d. h. jedes Element der Matrizenpotenzreihe) für $|\lambda| > \varrho$ konvergent ist; und zwar sind die Matrizen $\mathfrak{A}^{(m)}$ eindeutig bestimmt, da die Entwicklung (20) nur auf eine Weise möglich ist. Die Matrizen $\mathfrak{A}^{(m)}$ kann man im Anschluß an Liouville und C. Neumann sehr einfach berechnen. Zunächst ist das Gleichungssystem $(\lambda \mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{c}$ für $|\lambda| > \varrho$ auf genau eine Weise lösbar, und zwar wegen $(\lambda \mathfrak{E} - \mathfrak{A})^{-1} = \mathfrak{R}_\lambda$ durch $\mathfrak{x} = \mathfrak{R}_\lambda \mathfrak{c}$, so daß also für $|\lambda| > \varrho$

$$(21) \qquad \mathfrak{x} = \left(\sum_{m=0}^{\infty} \frac{\mathfrak{A}^{(m)}}{\lambda^{m+1}} \right) \mathfrak{c} = \sum_{m=0}^{\infty} \frac{\mathfrak{A}^{(m)} \mathfrak{c}}{\lambda^{m+1}}$$

gilt, und eine solche Entwicklung ist ebenfalls nur auf eine Weise möglich. Setzt man daher (21) in $(\lambda \mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{c}$ ein, so folgt durch

Vergleich der Koeffizienten von λ^{-m-1} die Rekursion

$$\mathfrak{E}(\mathfrak{A}^{(m+1)}\mathfrak{c}) - \mathfrak{A}(\mathfrak{A}^{(m)}\mathfrak{c}) = \mathfrak{o}; \qquad \mathfrak{E}(\mathfrak{A}^{(0)}\mathfrak{c}) = \mathfrak{c}$$

oder

$$\mathfrak{A}^{(m+1)}\mathfrak{c} = \mathfrak{A}(\mathfrak{A}^{(m)}\mathfrak{c}); \qquad \mathfrak{A}^{(0)}\mathfrak{c} = \mathfrak{c}.$$

Es ist daher $\mathfrak{A}^{(1)}\mathfrak{c} = \mathfrak{A}(\mathfrak{A}^{(0)}\mathfrak{c}) = \mathfrak{A}\mathfrak{c}$, mithin $\mathfrak{A}^{(2)}\mathfrak{c} = \mathfrak{A}(\mathfrak{A}^{(1)}\mathfrak{c}) = \mathfrak{A}(\mathfrak{A}\mathfrak{c}) = \mathfrak{A}^2\mathfrak{c}$ und allgemein $\mathfrak{A}^{(l)}\mathfrak{c} = \mathfrak{A}^l\mathfrak{c}$, rechterhand Matrizenpotenzierung verstanden. Dies muß für alle $\mathfrak{c}$, also insbesondere auch für die Einzelvektoren $\mathfrak{e}_i$ gelten, was jedoch, wie man sich leicht überzeugt, nur dann möglich ist, wenn $\mathfrak{A}^{(l)} = \mathfrak{A}^l$ gilt. Man hat daher

$$\mathfrak{R}_\lambda = \sum_{m=0}^{\infty} \frac{\mathfrak{A}^m}{\lambda^{m+1}} \quad \text{für} \quad |\lambda| > \varrho. \tag{22}$$

Wegen $\mathfrak{R}_\lambda = (\lambda\mathfrak{E} - \mathfrak{A})^{-1} = \frac{1}{\lambda\mathfrak{E} - \mathfrak{A}}$ hätten wir dies formal — ohne Konvergenzbeweis und dergleichen — auch kürzer herleiten können.

Da die Potenzreihen $R_{ik}(\lambda) = \sum_{m=0}^{\infty} (\lambda^{-1})^{m+1} a_{ik}^{(m)}$ für $|\lambda^{-1}| < \varrho^{-1}$ konvergent sind, so gibt es bekanntlich eine Schranke Ω derart, daß $|a_{ik}^{(m)}| < \Omega^m$ ist. Mithin sind die Reihen

$$\sum_{m=0}^{\infty} \frac{a_{ik}^{(m)} t^m}{m!}, \qquad \sum_{m=0}^{\infty} \frac{(-1)^m a_{ik}^{(2m)} t^{2m}}{(2m)!}, \qquad \sum_{m=0}^{\infty} \frac{(-1)^m a_{ik}^{(2m+1)} t^{2m+1}}{(2m+1)!}$$

für alle t konvergent. Wir können daher wegen $\|a_{ik}^{(m)}\| = \mathfrak{A}^m$ die Matrizen

$$e^{\mathfrak{A}t} \equiv \exp(\mathfrak{A}t) = \sum_{m=0}^{\infty} \frac{(\mathfrak{A}t)^m}{m!} = \left\| \sum_{m=0}^{\infty} \frac{a_{ik}^{(m)} t^m}{m!} \right\|, \tag{23}$$

$$\cos(\mathfrak{A}t) = \sum_{m=0}^{\infty} \frac{(-1)^m (\mathfrak{A}t)^{2m}}{(2m)!}, \qquad \sin(\mathfrak{A}t) = \sum_{m=0}^{\infty} \frac{(-1)^m (\mathfrak{A}t)^{2m+1}}{(2m+1)!}$$

einführen, die damit für alle $\mathfrak{A}$ und für alle t, also insbesondere auch für $t = 1$ erklärt sind. Es gelten, wie man sich leicht überzeugt, für diese Matrizentranszendenten im Falle *vertauschbarer* Argumente dieselben Funktionalgleichungen, wie für die entsprechenden gewöhnlichen elementaren Transzendenten, und zwar genau mit denselben Schlußweisen.

Mit Hilfe von $\exp(\mathfrak{A}t)$ kann man die allgemeine Lösung des Differentialsystems

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} a_{ik} x_k; \qquad i = 1, 2, \ldots, n \tag{24}$$

in derselben Gestalt darstellen, wie die allgemeine Lösung der einzigen Differentialgleichung $\dot{x} = a\,x$ durch e^{at}. Den Vektor mit den Komponenten $x_i(t)$ wollen wir durch $\mathfrak{x}_t$ bezeichnen. Das Differentialsystem lautet dann $\dot{\mathfrak{x}}_t = \mathfrak{A}\,\mathfrak{x}_t$, wobei für $t = 0$ der Anfangsvektor $\mathfrak{x}_0$ vorgeschrieben ist. Setzt man den Ansatz

$$(25)\qquad \mathfrak{x}_t = \sum_{m=0}^{\infty} \frac{\mathfrak{E}^{(m)}\,t^m}{m!}\,\mathfrak{x}_0$$

in die Vektordifferentialgleichung ein, so folgt durch Vergleich der Koeffizienten von t^m die Rekursion

$$(m+1)\frac{\mathfrak{E}^{(m+1)}}{(m+1)!}\mathfrak{x}_0 = \mathfrak{A}\left(\frac{\mathfrak{E}^{(m)}}{m!}\,\mathfrak{x}_0\right) = \frac{\mathfrak{A}\,\mathfrak{E}^{(m)}}{m!}\,\mathfrak{x}_0 \quad \text{mit} \quad \mathfrak{x}_0 = \mathfrak{E}^{(0)}\,\mathfrak{x}_0.$$

Da dies für alle $\mathfrak{x}_0$ gelten soll, muß $\mathfrak{E}^{(m+1)} = \mathfrak{A}\,\mathfrak{E}^{(m)}$, also wieder $\mathfrak{E}^{(m)} = \mathfrak{A}^m\,\mathfrak{E}^{(0)}$ sein. Nun ist aber $\mathfrak{x}_0 = \mathfrak{E}^{(0)}\,\mathfrak{x}_0$, also $\mathfrak{E}^{(0)} = \mathfrak{E}$, mithin $\mathfrak{E}^{(m)} = \mathfrak{A}^m$; $m = 0, 1, 2, \ldots$, so daß (25) in $\mathfrak{x}_t = [\exp(\mathfrak{A}\,t)]\,\mathfrak{x}_0$ übergeht. In Worten: Die Lösung $\mathfrak{x}_t$ entsteht aus dem Anfangsvektor $\mathfrak{x}_0$ durch Transformation mittels der Matrix $\exp(\mathfrak{A}\,t)$.

§ 11. Kanonische Matrizen. Die Schursche Transformation.

Wir wissen, daß die Determinante einer kanonischen Matrix $\|l_{ik}\|$ das Produkt der diagonalen Elemente l_{ii} ist (S. 11). Andererseits ist die Summe zweier kanonischen Matrizen wieder kanonisch. Insbesondere ist also die Matrix $\|\lambda\,\delta_{ik} - l_{ik}\| = \lambda\,\mathfrak{E} - \mathfrak{L}$ kanonisch, wenn $\mathfrak{L}$ es ist. Man hat somit in diesem Fall $\det(\lambda\,\mathfrak{E} - \mathfrak{L}) = (\lambda - l_{11})(\lambda - l_{22})\ldots(\lambda - l_{nn})$, so daß die Nullstellen des Säkularpolynoms, d. h. die Eigenwerte, einfach die diagonalen Elemente von $\mathfrak{L}$ sind. Da das Spektrum Matrizentransformationen gegenüber invariant ist, wird man daher versuchen, zu jeder Matrix $\mathfrak{A}$ eine nicht entartete Matrix $\mathfrak{S}$ derart zu finden, daß die transformierte Matrix $\mathfrak{L} = \mathfrak{S}\,\mathfrak{A}\,\mathfrak{S}^{-1}$ kanonisch ist. Bei der Resolvente $(\lambda\,\mathfrak{E} - \mathfrak{A})^{-1}$ verliert man durch die Transformation gar nichts. Denn $(\lambda\,\mathfrak{E} - \mathfrak{A})^{-1}$ transformiert sich nach S. 14 kogredient mit $(\lambda\,\mathfrak{E} - \mathfrak{A})$ und $(\lambda\,\mathfrak{E} - \mathfrak{A})$ transformiert sich nach § 8 kogredient mit $\mathfrak{A}$: ist $\mathfrak{R}_\lambda$ die Resolvente von $\mathfrak{A}$, so ist $\mathfrak{S}\,\mathfrak{R}_\lambda\,\mathfrak{S}^{-1}$ die Resolvente von $\mathfrak{S}\,\mathfrak{A}\,\mathfrak{S}^{-1}$.

Die Existenz einer kanonisierenden Transformation $\mathfrak{S}$ wird sogar unter der Einschränkung, daß $\mathfrak{S}$ unitär, d. h. $\mathfrak{S}^{-1} = \mathfrak{S}^*$ sein soll, durch die folgende überschlägliche Konstantenabzählung nahegelegt. $\mathfrak{A}$ hat n^2 gegebene Elemente; $\mathfrak{S}$ können wir beliebig wählen. Da aber $\mathfrak{S}\,\mathfrak{S}^* = \mathfrak{E}$ sein muß, haben wir in $\mathfrak{S} = \|s_{ik}\|$ nicht n^2, sondern nur $n^2 - (1 + 2 + \cdots + n)$ Parameter zur Verfügung. Denn damit

$\mathfrak{S} = \| s_{ik} \|$ unitär ist, haben die s_{ik} die Bedingungsgleichungen $\sum_{j=1}^{n} s_{ij} \bar{s}_{kj} = \delta_{ik}$ zu erfüllen, unter welchen es $\binom{n+1}{2}$ unabhängige gibt, so daß auch bei den unitären $\mathfrak{S}$ noch $\binom{n}{2} = n^2 - \binom{n+1}{2} = n^2 - (1+2+\cdots+n)$ Parameter verfügbar sind. Es ist daher zu erwarten, daß durch geschickte Wahl von $\mathfrak{S}$ ebenso viele Elemente von $\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$ zum Verschwinden gebracht werden können. Nun ist diese Anzahl eben mit der Anzahl der oberhalb der Diagonale stehenden Elemente identisch. Es ist zugleich die Vermutung nahegelegt, daß, wenn die n^2 Elemente von $\mathfrak{A}$ voneinander unabhängig sind, *mehr* als eine Kanonisierung, etwa eine Transformation auf die Diagonalgestalt, durch keinerlei Wahl von $\mathfrak{S}$ gelingen wird.

Wir beweisen nun den folgenden Satz von I. Schur: Zu jeder Matrix $\mathfrak{A}$ gibt es eine unitäre Matrix $\mathfrak{S} = \mathfrak{U}$ derart, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ kanonisch ist.

Es sei λ'_* ein Eigenwert der transponierten Matrix[1]) $\mathfrak{A}'$, so daß die zu $\lambda'_* \mathfrak{E} - \mathfrak{A}'$ gehörigen homogenen Gleichungen eine Eigenlösung haben, die aus Gründen der Homogenität normiert angenommen werden kann. Es sei also $|\mathfrak{r}_1|^2 = 1$ und $(\lambda'_* \mathfrak{E} - \mathfrak{A}')\mathfrak{r}_1 = \mathfrak{o}$, d. h. $\lambda'_* \mathfrak{r}_1 = \mathfrak{A}' \mathfrak{r}_1$. Wegen $\mathfrak{r}_1 \neq \mathfrak{o}$ können wir nach § 5 Vektoren $\mathfrak{r}_2, \ldots, \mathfrak{r}_n$ in der Anzahl $n-1$ derart bestimmen, daß die n Vektoren $\mathfrak{r}_i$ linear unabhängig sind. Mit Rücksicht auf den Schmidtschen Orthogonalisierungsprozeß (§ 6) können wir daher sogleich voraussetzen, daß $\mathfrak{r}_i \bar{\mathfrak{r}}_k = \delta_{ik}$ gilt für $i, k = 1, 2, \ldots, n$. Setzt man $\mathfrak{r}_i = (r_{i1}, r_{i2}, \ldots, r_{in})$, so besagt dies, daß die Matrix $\| r_{ik} \|$, die wir auch durch $\mathfrak{Q}_1$ bezeichnen wollen, unitär ist. Da $\mathfrak{Q}_1^{-1} = \bar{\mathfrak{Q}}'_1$ und $\lambda'_* \mathfrak{r}_1 = \mathfrak{A}' \mathfrak{r}_1$ sein muß, so ist das k-te Element der ersten Zeile der transformierten Matrix $\mathfrak{Q}_1 \mathfrak{A} \mathfrak{Q}_1^{-1} = \| g_{ik} \|$ gleich

$$g_{1k} = \sum_{j=1}^{n} \sum_{l=1}^{n} r_{1j} a_{jl} \bar{r}_{kl} = \sum_{l=1}^{n} \bar{r}_{kl} \sum_{j=1}^{n} a_{jl} r_{1j} = \bar{\mathfrak{r}}_k \cdot \mathfrak{A}' \mathfrak{r}_1 = \bar{\mathfrak{r}}_k \lambda'_* \mathfrak{r}_1 = \lambda'_* \delta_{1k},$$

somit für $k > 1$ gleich Null, d. h. es gilt

$$(26)\quad \mathfrak{Q}_1 \mathfrak{A} \mathfrak{Q}_1^{-1} = \begin{Vmatrix} \lambda'_* & 0 & 0 \ldots 0 \\ g_{21} & & \\ g_{31} & & \mathfrak{B} \\ \vdots & & \\ g_{n1} & & \end{Vmatrix}, \qquad \mathfrak{B} = \begin{Vmatrix} g_{22} & g_{23} & \cdots & g_{2n} \\ g_{32} & & & \\ \vdots & & & \\ g_{n2} & & & g_{nn} \end{Vmatrix},$$

---

[1]) Oder, was wegen $\det(\lambda\mathfrak{E} - \mathfrak{A}) = \det(\lambda\mathfrak{E} - \mathfrak{A})' = \det(\lambda\mathfrak{E} - \mathfrak{A}')$ dasselbe ist, ein Eigenwert von $\mathfrak{A}$ selbst.

wobei also die Matrix $\mathfrak{B}$ nur $n-1$ Zeilen und Kolonnen hat. Wir können nun $\mathfrak{B}$ demselben Verfahren wie vorher $\mathfrak{A}$ unterwerfen, also eine unitäre Matrix $\mathfrak{P}$ mit $n-1$ Zeilen und Kolonnen derart bestimmen, daß

$$(27)\qquad \mathfrak{P}\mathfrak{B}\mathfrak{P}^{-1} = \left\| \begin{matrix} h_{22} & 0 \; 0 \ldots 0 \\ h_{32} & \\ \vdots & \mathfrak{W} \\ h_{n2} & \end{matrix} \right\|, \qquad \mathfrak{W} = \left\| \begin{matrix} h_{33} & h_{34} \cdots h_{3n} \\ h_{43} & \\ \vdots & \\ h_{n3} & h_{nn} \end{matrix} \right\|$$

wird, wobei $\mathfrak{W}$ eine Matrix mit $n-2$ Zeilen und Kolonnen bezeichnet. Setzen wir nun

$$(28)\qquad \mathfrak{Q}_2 = \left\| \begin{matrix} 1 & 0 \; 0 \ldots 0 \\ 0 & \\ 0 & \mathfrak{P} \\ \vdots & \\ 0 & \end{matrix} \right\|,$$

so ist diese Matrix mit n Zeilen und Kolonnen offenbar ebenfalls unitär. Aus (26), (27), (28) folgt leicht, daß die Transformierte von $\mathfrak{Q}_1 \mathfrak{A} \mathfrak{Q}_1^{-1}$ durch $\mathfrak{Q}_2$ die Gestalt

$$\mathfrak{Q}_2 \mathfrak{Q}_1 \mathfrak{A} \mathfrak{Q}_1^{-1} \mathfrak{Q}_2^{-1} = \left\| \begin{matrix} j_{11} & 0 & 0 & 0 & \ldots & 0 \\ j_{21} & j_{22} & 0 & 0 & \ldots & 0 \\ j_{31} & j_{32} & j_{33} & j_{34} & \cdots & j_{3n} \\ \vdots & & & & & \\ j_{n1} & \ldots & \ldots & \ldots & \ldots & j_{nn} \end{matrix} \right\|$$

hat, so daß darin die oberhalb der Diagonale stehenden Elemente nicht nur, wie in (26), in der ersten Zeile verschwinden, sondern auch in der zweiten. So können wir weitergehen und die oberhalb der Diagonale liegenden Elemente von $\mathfrak{A}$ durch n-äre unitäre Transformationen sukzessive ausfegen. Nun bilden aber die unitären Matrizen eine Gruppe, womit alles bewiesen ist. — Des näheren handelt es sich dabei um folgendes. Zum Schluß haben wir $n-1$ unitäre Matrizen $\mathfrak{Q}_1, \mathfrak{Q}_2, \ldots, \mathfrak{Q}_{n-1}$ (alle mit n Zeilen und Kolonnen) derart, daß in der Matrix

$$(29)\qquad \mathfrak{Q}_{n-1} \mathfrak{Q}_{n-2} \cdots \mathfrak{Q}_2 \mathfrak{Q}_1 \mathfrak{A} \mathfrak{Q}_1^{-1} \mathfrak{Q}_2^{-1} \cdots \mathfrak{Q}_{n-2}^{-1} \mathfrak{Q}_{n-1}^{-1}$$

alle oberhalb der Diagonale stehenden Elemente verschwinden, so daß diese Matrix kanonisch ist. Nun kann für $\mathfrak{S}_2 (\mathfrak{S}_1 \mathfrak{A} \mathfrak{S}_1^{-1}) \mathfrak{S}_2^{-1}$ wegen $\mathfrak{S}_1^{-1} \mathfrak{S}_2^{-1} = (\mathfrak{S}_2 \mathfrak{S}_1)^{-1}$ auch $(\mathfrak{S}_2 \mathfrak{S}_1) \mathfrak{A} (\mathfrak{S}_2 \mathfrak{S}_1)^{-1}$ geschrieben werden. Ferner ist das Produkt zweier unitären Matrizen wieder unitär. Aus

$\mathfrak{S}_1^{-1} = \mathfrak{S}_1^*$ und $\mathfrak{S}_2^{-1} = \mathfrak{S}_2^*$ folgt nämlich $\mathfrak{S}_2^{-1}\mathfrak{S}_1^{-1} = \mathfrak{S}_2^*\mathfrak{S}_1^*$, also wegen $\mathfrak{S}_2^{-1}\mathfrak{S}_1^{-1} = (\mathfrak{S}_1\mathfrak{S}_2)^{-1}$ und $\mathfrak{S}_2^*\mathfrak{S}_1^* = (\mathfrak{S}_1\mathfrak{S}_2)^*$ auch $(\mathfrak{S}_1\mathfrak{S}_2)^{-1} = (\mathfrak{S}_1\mathfrak{S}_2)^*$. Setzt man daher $\mathfrak{U} = \mathfrak{Q}_{n-1}\mathfrak{Q}_{n-2}\cdots\mathfrak{Q}_2\mathfrak{Q}_1$, so ist $\mathfrak{U}$ unitär und die kanonische Matrix (29) gleich $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$. Damit ist der Schursche Satz bewiesen.

Der Schursche Satz und auch die Beweisanordnung stellt eine sinngemäße Verallgemeinerung des klassischen Verfahrens dar, das durch sukzessive orthogonale Abspaltung der Eigenwerte eine reell-symmetrische Matrix (eine quadratische Form) auf die Diagonalgestalt (in eine Summe von Quadraten) zu transformieren gestattet. Wir kommen darauf noch mehrfach zurück. — Bereits Lagrange und Gauß haben bemerkt, daß sich diese „Transformation auf die Hauptachsen" meistens viel bequemer gestaltet, wenn man den schrittweisen Abspaltungen nicht an die Matrix $\mathfrak{A}$ selbst, sondern an ihre Reziproke $\mathfrak{A}^{-1}$ anknüpft, deren Existenz zunächst vorausgesetzt wird. Jacobi hat diese, heute nach ihm benannte Transformation, die, sofern sie existiert, mit der Schurschen inhaltlich identisch ist, auf nicht-Hermitesche Matrizen übertragen. Allerdings gibt es Matrizen, wo die Jacobische Transformation, im Gegensatz zur Schurschen, versagt. Wir werden uns in §§ 27, 28 nur mit dem Spezialfall der Jacobischen Transformation beschäftigen, den wir späterhin, bei Untersuchungen über unendliche Matrizen, benötigen werden.

Aus dem Schurschen Satz folgt, daß $\det\mathfrak{A}$ immer gleich dem Produkt der Eigenwerte von $\mathfrak{A}$ ist[1]). Denn zunächst sind $\det\mathfrak{A}$ und die n Eigenwerte Matrizentransformationen gegenüber invariant, so daß nach dem Schurschen Satz die Matrix $\mathfrak{A}$ kanonisch angenommen werden kann. Nun ist aber die Determinante einer kanonischen Matrix das Produkt der diagonalen Elemente und diese sind bei kanonischen Matrizen mit den Eigenwerten identisch. — Ferner sind die Eigenwerte von $\mathfrak{A}'$ mit denjenigen von $\mathfrak{A}$ identisch (S. 21), so daß $\det\mathfrak{A} = \det\mathfrak{A}'$ als eine Selbstverständlichkeit erscheint. Ebenso läßt sich die Beziehung $\det(\overline{\mathfrak{A}}) = \overline{(\det\mathfrak{A})}$ dahin verschärfen, daß die Eigenwerte von $\mathfrak{A}$ zu denjenigen von $\overline{\mathfrak{A}}$ konjugiert sind. In der Tat ist $\det(\lambda\,\delta_{ik} - a_{ik}) = 0$ wegen (4) gleichbedeutend mit $\det(\bar{\lambda}\,\delta_{ik} - \bar{a}_{ik}) = 0$.

§ 12. Die normalen Matrizen.

Im Anschluß an den Schurschen Satz haben Schur und Toeplitz die Klasse derjenigen Matrizen $\mathfrak{A}$ untersucht, die mit ihrer begleiten-

---

[1]) Dies folgt auch unmittelbar aus dem Fundamentalsatz der Algebra.

den Matrix $\overline{\mathfrak{A}}' = \mathfrak{A}^*$ (deren Determinante nach dem vorhergehenden $= \overline{\det \mathfrak{A}}$ ist, während die Eigenwerte zu denjenigen von $\mathfrak{A}$ konjugiert sind) vertauschbar sind. Diese, der Bedingung $\mathfrak{A}\mathfrak{A}^* = \mathfrak{A}^*\mathfrak{A}$ genügenden Matrizen nennt Toeplitz *normal*. Der normale Charakter ist unitären Substitutionen gegenüber invariant. Aus $\mathfrak{A}\mathfrak{A}^* = \mathfrak{A}^*\mathfrak{A}$ und $\mathfrak{U}^{-1} = \mathfrak{U}^*$ folgt nämlich

$$(\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1})(\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1})^* = \mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}\mathfrak{U}^{-1*}\mathfrak{A}^*\mathfrak{U}^* = \mathfrak{U}\mathfrak{A}\mathfrak{A}^*\mathfrak{U}^* = \mathfrak{U}\mathfrak{A}^*\mathfrak{A}\mathfrak{U}^*$$
$$= (\mathfrak{U}^{-1*}\mathfrak{A}^*\mathfrak{U}^*)(\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}) = (\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1})^*(\mathfrak{U}\mathfrak{A}\mathfrak{A}^{-1}).$$

Die begleitende Matrix einer Diagonalmatrix ist wieder eine Diagonalmatrix und zwei Diagonalmatrizen sind gewiß vertauschbar, so daß die Diagonalmatrizen und daher auch die Matrizen, die unitär in Diagonalmatrizen transformiert werden können, gewiß normal sind. Damit ist aber die Klasse der normalen Matrizen auch erschöpft, indem jede normale Matrix unitär auf die Diagonalgestalt gebracht werden kann (so daß also die normalen Matrizen diejenigen Matrizen sind, für welche das „Hauptachsenproblem" gelöst werden kann). In Anbetracht des Schurschen Satzes beweisen wir mehr, wenn wir zeigen, daß eine normale Matrix, die kanonisch ist, sogar eine Diagonalmatrix sein muß.

Ist zunächst $\mathfrak{A}$ normal, so gilt $\sum_{j=1}^{n} a_{ij}\bar{a}_{kj} = \sum_{j=1}^{n} \bar{a}_{ji} a_{jk}$. Dies geht, wenn $\mathfrak{A}$ kanonisch, also $a_{ij} = 0$ für $i < j$ ist, mit $k = i$ in

$$\sum_{j=1}^{i} |a_{ij}|^2 = \sum_{j=i}^{n} |a_{ji}|^2, \tag{30}$$

also für $i = 1$ in $|a_{11}|^2 = \sum_{j=1}^{n} |a_{j1}|^2$ über, so daß $0 = \sum_{j=2}^{n} |a_{j1}|^2$ gilt, und daher alle a_{j1}; $j = 2, 3, \ldots, n$, d. h. alle unterhalb der Diagonale liegenden Elemente der ersten Kolonne gewiß verschwinden. Damit geht (30) für $i \geqq 2$ in

$$\sum_{j=2}^{i} |a_{ij}|^2 = \sum_{j=i}^{n} |a_{ji}|^2$$

über, woraus mit $i = 2$ wieder $|a_{22}|^2 = \sum_{j=2}^{n} |a_{j2}|^2$, d. h. $0 = \sum_{j=3}^{n} |a_{j2}|^2$ und daher das Verschwinden aller unterhalb der Diagonale liegenden Elemente der zweiten Kolonne folgt. Durch Wiederholung dieser Schlußweise ergibt sich, daß unterhalb der Diagonale lauter Nullen stehen, während dies oberhalb der Diagonale nach Voraussetzung der Fall ist, so daß $\|a_{ik}\|$, wie behauptet, eine Diagonalmatrix sein muß.

Die binäre Matrix

$$(31) \qquad \left\| \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix} \right\|$$

ist, wie man sich leicht überzeugt, mit ihrer Begleitenden nicht vertauschbar. Es gibt also Matrizen, die mit keiner Diagonalmatrix unitär äquivalent sind. Zwei Matrizen heißen dabei unitär äquivalent, wenn sie durch eine passend wählbare unitäre Transformation $\mathfrak{U}$ ineinander übergeführt werden können. Aus $\mathfrak{A} = \mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1}$ folgt $\mathfrak{U}^{-1}\mathfrak{A}\mathfrak{U} = \mathfrak{U}^{-1}\mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1}\mathfrak{U} = \mathfrak{B}$, d.h. $\mathfrak{B} = \mathfrak{T}\mathfrak{A}\mathfrak{T}^{-1}$, wobei $\mathfrak{T} = \mathfrak{U}^{-1}$, also unitär ist. Ferner ist die Eigenschaft eines Matrizenpaares $\mathfrak{A}, \mathfrak{B}$, unitär äquivalent zu sein, gegenüber kogredienten unitären Transformationen von $\mathfrak{A}$ und $\mathfrak{B}$ invariant. Ist nämlich $\mathfrak{A} = \mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1}$, wobei $\mathfrak{U}$ unitär ist, und bedeutet $\mathfrak{V}$ eine beliebige unitäre Matrix, so gibt es eine unitäre Matrix $\mathfrak{W}$ derart, daß $\mathfrak{V}\mathfrak{A}\mathfrak{V}^{-1} = \mathfrak{W}(\mathfrak{V}\mathfrak{B}\mathfrak{V}^{-1})\mathfrak{W}^{-1}$ wird. Zu diesem Ende muß $\mathfrak{W}$ nur derart gewählt werden, daß $(\mathfrak{V}\mathfrak{U})\mathfrak{B}(\mathfrak{V}\mathfrak{U})^{-1} = (\mathfrak{W}\mathfrak{V})\mathfrak{B}(\mathfrak{W}\mathfrak{V})^{-1}$ wird. Hierfür genügt es aber, wenn $\mathfrak{W}$ gemäß der Bedingung $\mathfrak{V}\mathfrak{U} = \mathfrak{W}\mathfrak{V}$ festgelegt, also $\mathfrak{W} = \mathfrak{V}\mathfrak{U}\mathfrak{V}^{-1}$ gesetzt wird. Dieses $\mathfrak{W}$ ist, als Produkt dreier unitären Matrizen, sicher unitär. Damit ist die unitäre Invarianz der unitären Äquivalenz bewiesen.

Wir wollen jetzt die Frage beantworten, wann zwei *normale* Matrizen $\mathfrak{A}, \mathfrak{B}$ unitär äquivalent sind. Hierfür ist wegen $\det(\lambda\mathfrak{E} - \mathfrak{A}) \equiv \det(\lambda\mathfrak{E} - \mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})$ gewiß notwendig, daß $\mathfrak{A}$ und $\mathfrak{B}$ dieselben Eigenwerte mit derselben bz. Vielfachheit, d. h. dasselbe Spektrum haben. Nun ist aber diese notwendige Bedingung, wie wir jetzt zeigen wollen, zugleich hinreichend, so daß also das Spektrum bei *normalen* Matrizen das *vollständige* unitäre Invariantensystem bildet.

Eine Matrix von n^2 Elementen möge (aus sogleich ersichtlichen Gründen) eine Permutationsmatrix genannt werden, wenn darin jede Zeile und jede Kolonne ein und nur ein Element enthält, das gleich $= 1$ ist, während die übrigen $n^2 - n$ Elemente verschwinden. Komponiert man die i-te Zeile einer Permutationsmatrix mit der k-ten, so ergibt sich δ_{ik}, so daß $\mathfrak{P}\mathfrak{P}' = \mathfrak{E}$ ist. Da $\mathfrak{P}$ reell ist, so kann man hierfür auch $\mathfrak{P}\overline{\mathfrak{P}}' = \mathfrak{E}$ schreiben, d. h. $\mathfrak{P}$ ist unitär. Dabei ist auch $\mathfrak{P}^{-1} = \mathfrak{P}'$ eine Permutationsmatrix, ferner ist das Produkt zweier Permutationsmatrizen, wie man leicht ausrechnet, eine Permutationsmatrix. Ist $\mathfrak{G} = \|\delta_{ik} g_i\|$ eine Diagonalmatrix, so ist es, wie man sich leicht überzeugt, auch $\mathfrak{P}\mathfrak{G}\mathfrak{P}^{-1} = \|\delta_{ik} h_i\|$, und zwar ist die Zahlenfolge $h_1, h_2, \ldots, h_n$ nur eine Permutation von $g_1, g_2, \ldots, g_n$, da das Spektrum, das bei Diagonalmatrizen durch die Elemente der Diagonale gebildet wird, invariant ist. Umgekehrt, sind die h_i nur eine Permu-

tation der g_i, so kann man auch eine Permutationsmatrix $\mathfrak{P}$ angeben derart, daß $\mathfrak{P}\,\|g_i\delta_{ik}\|\,\mathfrak{P}^{-1} = \|h_i\delta_{ik}\|$ wird. Der Beweis, der auf den Spezialfall $h_1 = g_2$, $h_2 = g_1$, $h_i = g_i$, $i > 2$ zurückgeführt werden kann, soll dem Leser überlassen bleiben.

Es seien nun $\mathfrak{A}$, $\mathfrak{B}$ zwei normale Matrizen mit demselben Spektrum. Es gibt dann zwei unitäre Matrizen $\mathfrak{U}$, $\mathfrak{V}$ derart, daß die beiden Matrizen $\mathfrak{G} = \mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$, $\mathfrak{H} = \mathfrak{V}\mathfrak{B}\mathfrak{V}^{-1}$ Diagonalmatrizen sind, deren Diagonalen auseinander durch eine Permutation entstehen, so daß es ein Permutationsmatrix $\mathfrak{P}$ mit $\mathfrak{H} = \mathfrak{P}\mathfrak{G}\mathfrak{P}^{-1}$ vorhanden ist. Man hat daher $\mathfrak{V}\mathfrak{B}\mathfrak{V}^{-1} = \mathfrak{P}\mathfrak{U}\,\mathfrak{A}\mathfrak{U}^{-1}\mathfrak{P}^{-1}$ oder $\mathfrak{B} = \mathfrak{V}^{-1}\mathfrak{P}\mathfrak{U}\,\mathfrak{A}\mathfrak{U}^{-1}\mathfrak{P}^{-1}\mathfrak{V}$, mithin $\mathfrak{B} = \overline{\mathfrak{W}}\mathfrak{A}\mathfrak{W}$, wenn $\mathfrak{W} = \mathfrak{V}^{-1}\mathfrak{P}\mathfrak{U}$ gesetzt wird. Nun ist $\mathfrak{W}$, als Produkt dreier unitären Matrizen, gewiß unitär. Mithin sind $\mathfrak{A}$ und $\mathfrak{B}$ unitär äquivalent, w. z. b. w.

§ 13. Die Resolvente der normalen Matrizen.

Die zur Matrix $\|\lambda\mathfrak{E} - \mathfrak{A}\|$ gehörigen inhomogenen Gleichungen sind, wenn $\mathfrak{A} = \|\delta_{ik}\lambda_i\|$ eine Diagonalmatrix ist, $\sum_{k=1}^{n}(\lambda\delta_{ik} - \delta_{ik}\lambda_i)\,x_k = c_i$, d. h. $(\lambda - \lambda_i)\,x_i = c_i$, so daß $x_i = \frac{c_i}{\lambda - \lambda_i}$ die einzige Lösung ist, wenn λ nicht ein Eigenwert λ_i wird. Andererseits ist die Resolvente $\|R_{ik}(\lambda)\|$ die Reziproke von $\lambda\mathfrak{E} - \mathfrak{A}$, also $x_i = \sum_{k=1}^{n} c_k R_{ik}(\lambda)$. Der Vergleich ergibt

$$(32) \qquad R_{ii}(\lambda) = \frac{1}{\lambda - \lambda_i}; \qquad R_{ik}(\lambda) \equiv 0, \qquad i \neq k.$$

Ist umgekehrt (32) erfüllt, so ist $\mathfrak{A}$ eine Diagonalmatrix. Wir benutzen dabei nur den Umstand, daß die Matrix $\|R_{ik}(\lambda)\|$ bei jedem λ eine Diagonalmatrix ist und schließen daraus, daß auch ihre Reziproke, $\|\lambda\mathfrak{E} - \mathfrak{A}\|$, und daher auch $\mathfrak{A}$ selbst, eine Diagonalmatrix sein muß. — Ist nun allgemeiner $\mathfrak{A}$ eine Matrix, bei welcher die $R_{ik}(\lambda)$ nur einfache Pole haben derart, daß in (18), S. 17 für $h > 1$ alle ϱ_{ik}^{hj} gewiß verschwinden, so werden die Elemente der Matrix $\mathfrak{S}\,\|R_{ik}(\lambda)\|\,\mathfrak{S}^{-1}$ ebenfalls nur einfache Pole haben. Aber auch nur dann, da auch $\mathfrak{S}^{-1}$ ein $\mathfrak{S}$ ist. Da $\mathfrak{S}\,\|R_{ik}(\lambda)\|\,\mathfrak{S}^{-1}$ die Resolvente von $\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$ ist, so ist daher die Eigenschaft von $\mathfrak{A}$, in der Resolvente nur einfache Pole zu haben, Matrizentransformationen gegenüber invariant. Da diese Eigenschaft wegen (32) allen Diagonalmatrizen zukommt, so kommt sie auch allen normalen Matrizen zu.

Es sei $\mathfrak{A}$ eine beliebige Matrix, $\lambda_1, \lambda_2, \ldots, \lambda_n$ seien ihre Eigenwerte, und $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(p)}$ ihre *verschiedenen* Eigenwerte, so daß jedes $\lambda^{(j)}$ unter

den λ_i genau s_j-mal vorkommt; vgl. (17). Es bezeichne ferner r_j die Anzahl ($\geqq 1$) der linear unabhängigen Lösungen, welche das zu $\lambda\mathfrak{E} - \mathfrak{A}$ gehörige homogene Gleichungssystem für $\lambda = \lambda^{(j)}$ zuläßt. Da die Zahlen $\lambda^{(j)}$, s_j, r_j, p, wie wir gesehen haben, Matrizentransformation gegenüber invariant sind, so dürfen wir annehmen, daß $\mathfrak{A}$ kanonisch ist. Es sei $\mathfrak{A}$ zunächst auch normal, also eine Diagonalmatrix. Die zu $\lambda\mathfrak{E} - \mathfrak{A}$ gehörigen homogenen Gleichungen sind dann $\lambda x_i - \lambda_i x_i = 0$, oder, wenn $\lambda = \lambda^{(j)}$ gesetzt wird, $(\lambda^{(j)} - \lambda_i) x_i = 0$, $i = 1, 2, \ldots, n$. Ist nun i derart, daß $\lambda^{(j)} \neq \lambda_i$ ist (j ist fest gewählt), so kann die zu diesem i gehörige Gleichung $(\lambda^{(j)} - \lambda_i) x_i = 0$ nur durch $x_i = 0$ erfüllt werden. Offenbar ist daher die Anzahl r_j der linear unabhängigen Lösungen des ganzen homogenen Systems identisch mit der Anzahl der linear unabhängigen Lösungen desjenigen Teilsystems $(\lambda^{(j)} - \lambda_i) x_i = 0$, in dem i nicht von 1 bis n geht, sondern nur die s_j ($\geqq 1$) Werte $i = i_{1j}, i_{2j}, \ldots, i_{s_j j}$ durchläuft, für welche $\lambda_i = \lambda^{(j)}$ ausfällt. Ist nun $i = i^{(0)}$ von dieser Beschaffenheit, so erhält man gewiß eine Eigenlösung, wenn man $x_{i^{(0)}} = 1$ und die übrigen $x_i = 0$ setzt. D. h., $\mathfrak{x} = \mathfrak{e}_{i_{\nu j}}$, wobei $\mathfrak{e}_k$ den k-ten Einzelvektor bezeichnet, ist gewiß eine Eigenlösung; $\nu = 1, 2, \ldots, s_j$, und offenbar kann jede Eigenlösung des Teilsystems aus diesen s_j Einzelvektoren, die, als *verschiedene* Einzelvektoren, linear unabhängig sind, komponiert werden, so daß s_j, die Häufigkeitszahl von $\lambda^{(j)}$ in der Folge $\lambda_1, \lambda_2, \ldots, \lambda_n$, d. h. die Vielfachheit des Eigenwertes $\lambda^{(j)}$, mit der Anzahl der zu $\lambda^{(j)} \cdot \mathfrak{E} - \mathfrak{A}$ gehörigen linear unabhängigen Lösungen übereinstimmt. Bei Diagonalmatrizen, also auch bei normalen Matrizen gilt daher

$$s_j = r_j; \quad j = 1, 2, \ldots, p.$$

[In der Weierstraßschen Terminologie besagt dies, daß die normalen Matrizen nur lineare Elementarteiler haben können[1]).]

Ist $\mathfrak{A} = \| a_{ik} \|$ eine Diagonalmatrix, so hat das Differentialsystem $\dot{\mathfrak{x}}_t = \mathfrak{A}\mathfrak{x}_t$ [vgl. (24)] die allgemeine Lösung $x_i = x_i(0) e^{\lambda_i t}$, $\lambda_i = a_{ii}$, es treten also keine „säkularen" Glieder auf, genauer die Zeit kommt nur im Exponenten vor, nicht aber auch als ein Faktorpolynom. Dies gilt offenbar bei allen Matrizen $\mathfrak{A}$, die auf die Diagonalform gebracht werden können, also auch bei allen normalen. Entsprechendes gilt bei den Systemen $\ddot{\mathfrak{x}}_t = \mathfrak{A}\mathfrak{x}$, usf. Vgl. übrigens S. 183 weiter unten.

---

[1]) Die Matrix $\mathfrak{A}$ hat dann und nur dann durchweg lineare Elementarteiler, wenn es eine nicht entartete, jedoch nicht notwendig unitäre Matrix $\mathfrak{S}$ gibt derart, daß $\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$ eine Diagonalmatrix ist. Dies ist dann und nur dann der Fall, wenn alle Singularitäten sämtlicher $R_{ik}(\lambda)$ einfache Pole sind, also gewiß dann (jedoch nicht nur dann), wenn die n Eigenwerte voneinander verschieden sind.

§ 14. Die beiden Normmatrizen. Hermitescher Charakter.

Es möge zunächst eine von Schur herrührende Charakterisierung der normalen Matrizen erwähnt werden, die u. a. den für die Fredholmsche Theorie grundlegenden Hadamardschen Determinantensatz enthält.

Es sei $\mathfrak{A} = \| a_{ik} \|$ eine beliebige Matrix. Die Summe $\sum\limits_{i=1}^{n} \sum\limits_{k=1}^{n} | a_{ik} |^2$ wollen wir aus Gründen, die sogleich ersichtlich sein werden, die Normalspur von $\mathfrak{A}$ nennen (nach Frobenius heißt sie „Spannung von $\mathfrak{A}$"). Die Normalspur ist unitären Transformationen gegenüber invariant [kann aber bei beliebigen Matrizen nicht mittels der Eigenwerte ausgedrückt werden, ist also eine unitäre, aber nicht spektrale Invariante]. Ist nämlich $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$, so ist die Normalspur von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ gleich

$$\sum_{i=1}^{n} \sum_{k=1}^{n} \Big| \sum_{j=1}^{n} \sum_{l=1}^{n} u_{ij} a_{jl} \bar{u}_{kl} \Big|^2 = \sum_{i=1}^{n} \sum_{k=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{n} u_{ij} a_{jl} \bar{u}_{kl} \sum_{g=1}^{n} \sum_{h=1}^{n} \bar{u}_{ig} \bar{a}_{gh} u_{kh}$$

$$= \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{g=1}^{n} \sum_{h=1}^{n} a_{jl} \bar{a}_{gh} \sum_{i=1}^{n} u_{ij} \bar{u}_{ig} \sum_{k=1}^{n} \bar{u}_{kl} u_{kh}$$

$$= \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{g=1}^{n} \sum_{h=1}^{n} a_{jl} \bar{a}_{gh} \delta_{jg} \delta_{lh} = \sum_{j=1}^{n} \sum_{l=1}^{n} a_{jl} \sum_{g=1}^{n} \delta_{jg} \sum_{h=1}^{n} \bar{a}_{gh} \delta_{lh}$$

$$= \sum_{j=1}^{n} \sum_{l=1}^{n} a_{jl} \sum_{g=1}^{n} \delta_{jg} \bar{a}_{gl} = \sum_{j=1}^{n} \sum_{l=1}^{n} a_{jl} \bar{a}_{jl} = \sum_{i=1}^{n} \sum_{k=1}^{n} | a_{ik} |^2 .$$

Wählen wir nun $\mathfrak{U}$ so, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ kanonisch ist. $\mathfrak{A}$ ist dann und nur dann normal, wenn $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ zugleich diagonal ist. In diesem Falle ist die Normalspur von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ gleich der Quadratsumme der Beträge der Diagonalelemente, d. h. der Eigenwerte von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$, und sonst größer. Nun haben $\mathfrak{A}$ und $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ dieselben Eigenwerte λ_ν und dieselbe Normalspur. Es ist also bei jeder Matrix $\mathfrak{A}$

$$\sum_{\nu=1}^{n} | \lambda_\nu |^2 \leqq \sum_{i=1}^{n} \sum_{k=1}^{n} | a_{ik} |^2, \tag{33}$$

und das Gleichheitszeichen gilt dann und nur dann, wenn $\mathfrak{A}$ normal ist.

Die Matrix $\mathfrak{A}^* \mathfrak{A}$ heißt die vordere, die Matrix $\mathfrak{A}\mathfrak{A}^*$ die hintere Norm von $\mathfrak{A}$; die beiden Normen sind dann und nur dann gleich, wenn $\mathfrak{A}$ mit $\mathfrak{A}^*$ vertauschbar, d. h. normal ist. Die Normalspur der vorderen Norm von $\mathfrak{A}$ ist auch bei nicht normalen Matrizen gleich

$$\sum_{i=1}^{n} \sum_{k=1}^{n} \Big| \sum_{j=1}^{n} \bar{a}_{ji} a_{jk} \Big|^2 = \sum_{i=1}^{n} \sum_{k=1}^{n} \Big(\sum_{j=1}^{n} \bar{a}_{ji} a_{jk} \Big) \Big(\sum_{l=1}^{n} a_{li} \bar{a}_{lk} \Big)$$

$$= \sum_{j=1}^{n} \sum_{l=1}^{n} \Big(\sum_{k=1}^{n} a_{jk} \bar{a}_{lk} \Big) \Big(\sum_{i=1}^{n} \bar{a}_{ji} a_{li} \Big) = \sum_{j=1}^{n} \sum_{l=1}^{n} \Big| \sum_{k=1}^{n} a_{jk} \bar{a}_{lk} \Big|^2 ,$$

also gleich der Normalspur der hinteren Norm von $\mathfrak{A}$.

Spur einer beliebigen Matrix $\mathfrak{A}$ heißt die Summe der Eigenwerte von $\mathfrak{A}$. Die Spur der Normen einer *normalen* Matrix $\mathfrak{A}$ ist also gleich der Normalspur von $\mathfrak{A}$. Es folgt nämlich etwa nach einer unitären Transformation von $\mathfrak{A}$ auf die Diagonalgestalt, daß die Eigenwerte der Norm von $\mathfrak{A}$ die Quadrate der absoluten Beträge der Eigenwerte von $\mathfrak{A}$ sind.

Die Matrix $\mathfrak{A}$ heißt von Hermiteschem Typus, wenn sie mit ihrer begleitenden Matrix $\overline{\mathfrak{A}}' = \mathfrak{A}^*$ identisch ist, so daß für alle Zeiger $a_{ik} = \bar{a}_{ki}$ gilt. Die diagonalen Elemente sind dann also gewiß reell. Da mit sich selbst jede Matrix vertauschbar ist, so sind die Hermiteschen Matrizen gewiß normal. Nicht jede normale Matrix ist Hermitesch. Hingegen sind die beiden Normen jeder, auch nicht normalen Matrix $\mathfrak{C}$, da $(\mathfrak{C}^* \mathfrak{C})^* = \mathfrak{C}^* \mathfrak{C}^{**} = \mathfrak{C}^* \mathfrak{C}$ gilt und die hintere Norm von $\mathfrak{C}$ die vordere Norm von $\mathfrak{C}^*$ ist, gewiß Hermitesch, und, wie wir bei (51) sehen werden [vgl. S. 30], „nichtnegativ definit"; und zwar sind beide Normen „positiv definit" oder keine der beiden ist es, je nachdem $\mathfrak{C}^{-1}$ existiert oder aber $\mathfrak{C}$ entartet ist (die Definitionen siehe im nächstfolgenden §). Es ist also nicht möglich, daß eine der beiden Normen positiv definit ist und die andere nicht. Bei *unendlichen* Matrizen wird dies allerdings vorkommen können (siehe S. 138).

Wir wollen nun zeigen, daß die beiden Normen *jeder* Matrix unitär äquivalent sind, so daß es zu jedem $\mathfrak{A}$ ein $\mathfrak{U}$ gibt derart, daß $\mathfrak{A}\mathfrak{A}^* = \mathfrak{U}(\mathfrak{A}^* \mathfrak{A})\mathfrak{U}^{-1}$ wird (ist $\mathfrak{A}$ normal, so kann man einfach $\mathfrak{U} = \mathfrak{E}$ setzen und umgekehrt). Da die beiden Normen stets normal sind, so läuft die Behauptung darauf hinaus, daß die beiden Normen dieselben Eigenwerte oder dasselbe Säkularpolynom haben. Denn die Übereinstimmung der Spektra ist nach § 12 die notwendige und hinreichende Bedingung für die unitäre Äquivalenz zweier normalen Matrizen. — Es sei zunächst $\mathfrak{B}$ eine Matrix derart, daß $\mathfrak{B}^{-1}$ existiert. Wegen $\mathfrak{B}^{-1}\mathfrak{B}\mathfrak{B}^*\mathfrak{B} = \mathfrak{B}^*\mathfrak{B}$ gibt es dann gewiß eine Matrix $\mathfrak{S}$ derart, daß $\mathfrak{S}\mathfrak{B}\mathfrak{B}^*\mathfrak{S}^{-1} = \mathfrak{B}^*\mathfrak{B}$ wird: $\mathfrak{S} = \mathfrak{B}^{-1}$. Nach § 8 folgt daraus

$$\det(\lambda\,\mathfrak{E} - \mathfrak{B}\mathfrak{B}^*) \equiv \det(\lambda\,\mathfrak{E} - \mathfrak{B}^*\mathfrak{B}).$$

Es sei nun $\mathfrak{A}$ eine beliebige Matrix. Den Fall, wo $\lambda = 0$ kein Eigenwert ist, so daß $\mathfrak{A}^{-1}$ existiert, haben wir soeben erledigt. Es sei also $\lambda = 0$ eine Nullstelle des Polynoms $\det(\lambda\,\mathfrak{E} - \mathfrak{A})$. Da dieses nur endlich viele Nullstellen hat, so gibt es ein $\eta > 0$ derart, daß $\det(\varepsilon\,\mathfrak{E} - \mathfrak{A}) \neq 0$ ist für $0 < |\varepsilon| < \eta$, so daß dabei $(\varepsilon\,\mathfrak{E} - \mathfrak{A})^{-1}$ existiert, und daher $\mathfrak{B}$ in der soeben hergeleiteten Identität gleich $\varepsilon\,\mathfrak{E} - \mathfrak{A}$

gesetzt werden darf. Nimmt man sodann in den beiderseits stehenden Polynomen den Grenzübergang $\varepsilon \to 0$ vor, so bleibt offenbar

$$\det(\lambda\mathfrak{E} - \mathfrak{A}\mathfrak{A}^*) \equiv \det(\lambda\mathfrak{E} - \mathfrak{A}^*\mathfrak{A}),$$

w. z. b. w.

§ 15. Hermitesche Matrizen.

Die beiden Normen einer Hermiteschen Matrix sind $= \mathfrak{A}^2$. Die Eigenschaft einer Matrix, von Hermiteschem Typus zu sein, ist unitären Transformationen gegenüber invariant. Aus $\mathfrak{A} = \mathfrak{A}^*$ und $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$ folgt nämlich $(\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1})^* = \mathfrak{U}^{-1*}\mathfrak{A}^*\mathfrak{U}^* = \mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$. Wählt man $\mathfrak{U}$ derart, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ eine Diagonalmatrix ist, so folgt daraus, daß die diagonalen Elemente von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$, d. h. die Eigenwerte von $\mathfrak{A}$, reell sein müssen. Sind um gekehrt die Eigenwerte einer *normalen* Matrix durchweg reell, so ist sie offenbar mit einer reellen Diagonalmatrix unitär äquivalent. Nun sind die reellen (und nur die reellen) Diagonalmatrizen Hermitesch, da bei Diagonalmatrizen die Forderung $a_{ik} = \bar{a}_{ki}$ wegen $0 = \bar{0}$ auf die Realität der a_{ii} hinausläuft. Mithin sind die Hermiteschen Matrizen diejenigen normalen Matrizen, deren Eigenwerte, die auch mehrfach sein können, sämtlich reell sind. Da die Eigenwerte von $\mathfrak{A}^2$ bei jedem $\mathfrak{A}$ die Quadrate der Eigenwerte von $\mathfrak{A}$ sind, so sind die Eigenwerte von $\mathfrak{A}^2$ gewiß $\geqq 0$, wenn $\mathfrak{A}$ Hermitesch ist. Eine Hermitesche Matrix, bei der alle Eigenwerte $\geqq 0$ sind, nennt man nichtnegativ definit. Eine nichtnegativ definite Hermitesche Matrix, deren Eigenwerte sogar sämtlich > 0 sind, heißt positiv definit. Das Quadrat einer Hermiteschen Matrix ist wegen $(\mathfrak{A}^2)^* = \mathfrak{A}^*\mathfrak{A}^* = \mathfrak{A}\mathfrak{A} = \mathfrak{A}^2$ gewiß Hermitesch. Das Quadrat einer Hermiteschen Matrix ist also eine nichtnegativ definite Hermitesche Matrix. Umgekehrt gibt es zu jeder nichtnegativ definiten Hermiteschen Matrix $\mathfrak{A}$ Hermitesche Matrizen $\mathfrak{B}$ derart, daß $\mathfrak{A} = \mathfrak{B}^2$ wird. Wegen $\mathfrak{U}\mathfrak{C}^2\mathfrak{U}^{-1} = (\mathfrak{U}\mathfrak{C}\mathfrak{U}^{-1})^2$ genügt es, dies unter der Voraussetzung zu zeigen, daß $\mathfrak{A} = \| a_{ii}\delta_{ik} \|$, also eine Diagonalmatrix mit $a_{ii} \geqq 0$ ist. Dann wird aber die Gleichung $\mathfrak{A} = \mathfrak{B}^2$ durch die wegen $\sqrt{a_{ii}} \gtreqless 0$ gewiß Hermiteschen Diagonalmatrix $\mathfrak{B} = \| \sqrt{a_{ii}}\,\delta_{ik} \|$ offenbar befriedigt. Die nichtnegativ definiten Hermiteschen Matrizen sind also diejenigen Matrizen, die als Quadrate von Hermiteschen (oder auch von nichtnegativ definiten Hermiteschen) Matrizen dargestellt werden können, während die positiv definiten Hermiteschen Matrizen und nur diese offenbar Quadrate sogar von positiv definiten Matrizen sind. Da eine Matrix $\mathfrak{A}$ dann und nur dann entartet ist $(\det\mathfrak{A} = 0)$, wenn mindestens ein Eigenwert verschwindet, und da die Quadrate der

Eigenwerte von $\mathfrak{A}$ die Eigenwerte von $\mathfrak{A}^2$ sind, so hat eine Hermitesche Matrix $\mathfrak{A}$ dann und nur dann eine Reziproke, wenn $\mathfrak{A}^2$ (nicht nur nichtnegativ definit, sondern sogar) positiv definit ist.

Ist $\mathfrak{A}$ Hermitesch und nicht entartet, so ist auch $\mathfrak{A}^{-1}$ Hermitesch. Aus $\mathfrak{A}\mathfrak{B} = \mathfrak{E}$ und $\mathfrak{A} = \mathfrak{A}^*$ folgt nämlich $\mathfrak{B}^*\mathfrak{A}^* = \mathfrak{B}^*\mathfrak{A} = \mathfrak{E}$, also $\mathfrak{B}^* = \mathfrak{B}$, da $\mathfrak{A}$ nur eine Reziproke haben kann. Ist außerdem $\mathfrak{A}$ positiv definit, so ist es auch $\mathfrak{A}^{-1}$. Denn die Eigenwerte von $\mathfrak{A}^{-1}$ sind die reziproken Eigenwerte von $\mathfrak{A}$, also ebenfalls positiv.

Das Produkt zweier Hermiteschen Matrizen ist dann und nur dann Hermitesch, wenn sie vertauschbar sind. Ist nämlich $\mathfrak{A} = \mathfrak{A}^*$ und $\mathfrak{B} = \mathfrak{B}^*$, so folgt aus $(\mathfrak{A}\mathfrak{B})^* = \mathfrak{A}\mathfrak{B}$ wegen $(\mathfrak{A}\mathfrak{B})^* = \mathfrak{B}^*\mathfrak{A}^*$ offenbar $\mathfrak{B}\mathfrak{A} = \mathfrak{A}\mathfrak{B}$ und umgekehrt. Es läßt sich übrigens ohne Mühe zeigen, daß zwei Hermitesche Matrizen dann und nur dann vertauschbar sind, wenn sie *kogredient* auf die Hauptachsen gedreht werden können, d. h. wenn es eine *gemeinsame* unitäre Matrix $\mathfrak{U}$ gibt derart, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ und $\mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1}$ Diagonalmatrizen sind[1]).

§ 16. Hauptachsentransformation.

Es soll jetzt der Zusammenhang zwischen den Eigenwerten und Eigenlösungen einerseits und der unitären Transformation auf die Diagonalgestalt (geometrisch: Hauptachsentransformation) andererseits besprochen werden. In Verbindung mit in § 13 berührten Fragen ist die Theorie des Elementarteiles daraus entstanden, und die Schursche Transformation stellt seine Übertragung auf beliebige Matrizen dar.

Sind zunächst $\mathfrak{x} = (x_1, x_2, \ldots, x_n)$, $\mathfrak{y} = (y_1, y_2, \ldots, y_n)$ zu zwei verschiedenen Eigenwerten ν, μ gehörige Lösungen des homogenen Systems $(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{z} = \mathfrak{o}$, wobei $\mathfrak{A}$ Hermitesch ist, so sind die beiden Vektoren $\mathfrak{x}, \mathfrak{y}$ orthogonal. Aus $\nu\mathfrak{y} = \mathfrak{A}\mathfrak{y}$ folgt nämlich, da ν reell ist, $\nu\bar{\mathfrak{y}} = \overline{\mathfrak{A}}\bar{\mathfrak{y}}$, also $\nu\mathfrak{x}\bar{\mathfrak{y}} = \mathfrak{x}(\overline{\mathfrak{A}}\bar{\mathfrak{y}})$, und aus $\mu\mathfrak{x} = \mathfrak{A}\mathfrak{x}$ folgt wegen $\mathfrak{a}\mathfrak{b} = \mathfrak{b}\mathfrak{a}$ entsprechend $\mu\mathfrak{x}\bar{\mathfrak{y}} = \bar{\mathfrak{y}}(\mathfrak{A}\mathfrak{x})$, mithin $(\nu - \mu)\mathfrak{x}\bar{\mathfrak{y}} = \mathfrak{x}(\overline{\mathfrak{A}}\bar{\mathfrak{y}}) - \bar{\mathfrak{y}}(\mathfrak{A}\mathfrak{x})$, so daß die Behauptung, $\mathfrak{x}\mathfrak{y} = 0$, gleichwertig ist mit $\mathfrak{x}(\overline{\mathfrak{A}}\bar{\mathfrak{y}}) = \bar{\mathfrak{y}}(\mathfrak{A}\mathfrak{x})$, d. h. $\sum_{i=1}^{n} x_i \sum_{k=1}^{n} \bar{a}_{ik}\bar{y}_k = \sum_{k=1}^{n} \bar{y}_k \sum_{i=1}^{n} a_{ki}x_i$. Dies ist aber wegen $a_{ik} = \bar{a}_{ki}$ eine richtige Gleichung.

---

[1]) Allgemeiner kann eine (endliche oder unendliche) Gesamtheit Ω von n-ären Hermiteschen Matrizen dann und nur mittels derselben unitären Transformation auf die Diagonalgestalt gebracht werden, wenn die Matrizen von Ω paarweise vertauschbar sind. — Entsprechendes gilt für unitäre Matrizen. Merkwürdige analytische Anwendungen gab davon H. Weyl. Vgl. z. B. H. Weyl, Math. Annalen **97** (1927), S. 338ff.; F. Peter und H. Weyl, ibid. S. 737ff.

Ist λ ein s-facher Eigenwert, so haben die zu $\lambda\mathfrak{E}-\mathfrak{A}$ gehörigen homogenen Gleichungen nach § 13, da $\mathfrak{A}$ Hermitesch, also gewiß normal ist, $s\,(\leqq n)$ linear-unabhängige Lösungen $\mathfrak{x}_1, \mathfrak{x}_2, \ldots, \mathfrak{x}_s$. Man zeigt nach dem Schmidtschen Orthogonalisationsverfahren, daß sie normiert und orthogonal angenommen werden können.

Es seien nun $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(p)}$ die verschiedenen Eigenwerte von $\mathfrak{A}$, es sei s_j die Vielfachheit von $\lambda^{(j)}$ und es seien $\mathfrak{x}_1^{(j)}, \mathfrak{x}_2^{(j)}, \ldots, \mathfrak{x}_{s_j}^{(j)}$ irgendwie gewählte, bei festem j normiert orthogonale, zu $\lambda^{(j)}$ gehörige Eigenlösungen. — Zwei $\mathfrak{x}$, die zu zwei verschiedenen j, also zu zwei verschiedenen λ gehören, sind, wie wir soeben gesehen haben, gewiß orthogonal. Wir haben also $s_1+s_2+\cdots+s_p=n$ Vektoren $\mathfrak{x}$, die normiert und paarweise orthogonal sind. Wir wollen sie nunmehr fortlaufend numerieren, $\mathfrak{x}_1, \mathfrak{x}_2, \ldots, \mathfrak{x}_n$, und die zugehörigen, voneinander nicht notwendig verschiedenen Eigenwerte entsprechend mit $\lambda_1, \lambda_2, \ldots, \lambda_n$ bezeichnen. Dann ist

$$(34)\qquad \lambda_i\mathfrak{x}_i=\mathfrak{A}\mathfrak{x}_i,\qquad \mathfrak{x}_i\bar{\mathfrak{x}}_k=\delta_{ik};\qquad i,k=1,2,\ldots,n.$$

Setzen wir $\mathfrak{x}_k=(\bar{u}_{k1}, \bar{u}_{k2}, \ldots, \bar{u}_{kn})$, so kann $\mathfrak{x}_i\bar{\mathfrak{x}}_k=\delta_{ik}$ dahin gedeutet werden, daß die Matrix $\mathfrak{U}=\|u_{ik}\|$ unitär ist. Nun ist diese, aus den Eigenlösungen hergeleitete Matrix dem $\mathfrak{A}$ derart angepaßt, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ diagonal ist. Aus (34) folgt nämlich $\lambda_k\bar{u}_{kj}=\sum_{l=1}^{n}a_{jl}\bar{u}_{kl}$, also wegen $\mathfrak{U}^{-1}=\mathfrak{U}^*$ auch

$$\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}=\Big\|\sum_{j=1}^{n}u_{ij}\sum_{l=1}^{n}a_{jl}\bar{u}_{kl}\Big\|=\Big\|\sum_{j=1}^{n}u_{ij}\lambda_k\bar{u}_{kj}\Big\|=\|\lambda_k\bar{\mathfrak{x}}_i\mathfrak{x}_k\|=\|\lambda_k\delta_{ik}\|.$$

§ 17. Hermitesche Komponenten. Die Kopplungsform.

Wir kehren nunmehr zu beliebigen (nicht notwendig normalen) Matrizen zurück und beweisen zunächst nach Toeplitz, daß jede Matrix $\mathfrak{A}=\|a_{ik}\|$ auf genau eine Weise in der Gestalt $\mathfrak{A}_{(1)}+\sqrt{-1}\,\mathfrak{A}_{(2)}$ geschrieben werden kann, wobei $\mathfrak{A}_{(1)}$ und $\mathfrak{A}_{(2)}$ von Hermiteschem Typus sind (insbesondere ist also die Nullmatrix die einzige Matrix, die mit $\sqrt{-1}$ multipliziert Hermitesch bleibt). Setzt man zunächst $\mathfrak{A}_{(1)}=\frac{1}{2}\mathfrak{A}+\frac{1}{2}\mathfrak{A}^*$, $\mathfrak{A}_{(2)}=\frac{1}{2}\sqrt{-1}\,\mathfrak{A}^*-\frac{1}{2}\sqrt{-1}\,\mathfrak{A}$, so ist $\mathfrak{A}=\mathfrak{A}_{(1)}+\sqrt{-1}\,\mathfrak{A}_{(2)}$ gewiß erfüllt, und $\mathfrak{A}_{(1)}$ hat aus ersichtlichen Symmetriegründen, $\mathfrak{A}_{(2)}$ aber wegen

$$\mathfrak{A}_{(2)}^*=\left\|\sqrt{-1}\,\frac{\bar{a}_{ki}-a_{ik}}{2}\right\|^*=\left\|\sqrt{-1}\,\frac{\bar{\bar{a}}_{ki}-\bar{a}_{ik}}{2}\right\|'=\left\|-\sqrt{-1}\,\frac{a_{ki}-\bar{a}_{ik}}{2}\right\|'$$

$$=\left\|-\sqrt{-1}\,\frac{a_{ik}-\bar{a}_{ki}}{2}\right\|=\mathfrak{A}_{(2)}$$

den Hermiteschen Charakter, $\mathfrak{H} = \overline{\mathfrak{H}}' = \mathfrak{H}^*$. Um die Eindeutigkeit dieser Zerlegung in Hermitesche Komponenten nachzuweisen, nehmen wir an, daß

$$\text{(I)} \qquad \mathfrak{A} = \mathfrak{A}_{(1)} + \sqrt{-1}\,\mathfrak{A}_{(2)}, \quad \mathfrak{A}_{(1)} = \mathfrak{A}_{(1)}^*, \quad \mathfrak{A}_{(2)} = \mathfrak{A}_{(2)}^*$$

und zugleich

$$\text{(II)} \qquad \mathfrak{A} = \mathfrak{B}_{(1)} + \sqrt{-1}\,\mathfrak{B}_{(2)}, \quad \mathfrak{B}_{(1)} = \mathfrak{B}_{(1)}^*, \quad \mathfrak{B}_{(2)} = \mathfrak{B}_{(2)}^*$$

gilt. Wegen (I) ist $\mathfrak{A}^* = (\mathfrak{A}_{(1)} + \sqrt{-1}\,\mathfrak{A}_{(2)})^* = \mathfrak{A}_{(1)}^* - \sqrt{-1}\,\mathfrak{A}_{(2)}^*$, also $\mathfrak{A}^* = \mathfrak{A}_{(1)} - \sqrt{-1}\,\mathfrak{A}_{(2)}$. Aus (II) folgt ebenso $\mathfrak{A}^* = \mathfrak{B}_{(1)} - \sqrt{-1}\,\mathfrak{B}_{(2)}$. Man hat daher

$$\text{(III)} \qquad \mathfrak{A}_{(1)} - \sqrt{-1}\,\mathfrak{A}_{(2)} = \mathfrak{B}_{(1)} - \sqrt{-1}\,\mathfrak{B}_{(2)},$$

während wegen (I), (II) auch

$$\text{(IV)} \qquad \mathfrak{A}_{(1)} + \sqrt{-1}\,\mathfrak{A}_{(2)} = \mathfrak{B}_{(1)} + \sqrt{-1}\,\mathfrak{B}_{(2)}$$

gilt. Indem (III) mit ± 1 multipliziert zu (IV) addiert wird, ergibt sich $\mathfrak{A}_{(1)} = \mathfrak{B}_{(1)}$, $\mathfrak{A}_{(2)} = \mathfrak{B}_{(2)}$, w. z. b. w.

Man überzeugt sich leicht, daß aus $\mathfrak{A}\mathfrak{A}^* = \mathfrak{A}^*\mathfrak{A}$ stets $\mathfrak{A}_{(1)}\mathfrak{A}_{(2)} = \mathfrak{A}_{(2)}\mathfrak{A}_{(1)}$ folgt und umgekehrt. M. a. W., eine Matrix $\mathfrak{A}$ ist dann und nur dann normal, wenn ihre beiden Hermiteschen Komponenten vertauschbar sind. Diese Bemerkung scheint für eine Spektraltheorie der unendlichen normalen (d. h. mit $\mathfrak{A}^*$ vertauschbaren) Matrizen $\mathfrak{A}$ von Wichtigkeit zu sein. — Hierher gehört auch die folgende Bemerkung von Frobenius: Ist $\mathfrak{A}$ eine normale Matrix, die eine Reziproke hat, so ist die Matrix $\mathfrak{A}^*\mathfrak{A}^{-1}$ unitär.

Es bezeichne E die „komplexe Einheitskugelfläche der Variablen $x_1, x_2, \ldots, x_n$", d. i. die Mannigfaltigkeit aller, der Bedingung

$$\text{(35)} \qquad \mathsf{E}: \quad \sum_{i=1}^{n} x_i \bar{x}_i = \sum_{i=1}^{n} |x_i|^2 = 1$$

genügenden, reellen oder komplexen Zahlensysteme $(x_1, x_2, \ldots, x_n)$, oder die Gesamtheit aller normierten Vektoren $\mathfrak{x}$. Hierbei ist n, wie bisher, eine irgendwie, jedoch ein für allemal gewählte natürliche Zahl. Wir ordnen jeder Matrix $\mathfrak{A}$ als „Kopplungsform" die folgende Funktion der n Variablen zu:

$$\text{(36)} \qquad \Phi(\mathfrak{A}) = \Phi(\mathfrak{A}; \mathfrak{x}) = \Phi(\mathfrak{A}; x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} x_i \bar{x}_k.$$

Bewegt sich $\mathfrak{x}$ auf E, so beschreibt $\Phi(\mathfrak{A}; \mathfrak{x})$ in der komplexen Φ-Ebene eine Punktmenge, die mit $\mathsf{W}(\mathfrak{A})$ bezeichnet werden soll. Die

Gesamtheit der Werte, derer $\Phi(\mathfrak{A})$ auf E fähig ist, heißt der Wertevorrat von $\mathfrak{A}$; er wird durch $\mathsf{W}(\mathfrak{A})$ repräsentiert. Die obere Grenze der Beträge der zu dem Wertevorrat von $\mathfrak{A}$ gehörigen Zahlen bezeichnen wir mit $\mathsf{M}(\mathfrak{A})$, so daß $\mathsf{M}(\mathfrak{A})$ die kleinste Zahl bedeutet, für welche $|\Phi(\mathfrak{A};\mathfrak{x})| \leqq \mathsf{M}(\mathfrak{A})$ in jedem Punkte von E gilt, oder auch den Radius des kleinsten, um den Nullpunkt der Φ-Ebene geschlagenen Kreisbereiches, der alle Punkte von $\mathsf{W}(\mathfrak{A})$ enthält. Übrigens ist die obere Grenze zugleich ein Maximum. Denn Φ ist stetig und E (also auch $\mathsf{W}(\mathfrak{A})$) abgeschlossen. Doch wollen wir dies nach Möglichkeit nicht ausnutzen, damit die Sprechweise auch bei den später zu betrachtenden unendlichen Matrizen in Kraft bleibt.

Unterwirft man einen Punkt $\mathfrak{x}$ von E einer unitären Transformation $\mathfrak{U}$, so liegt der Bildpunkt $\mathfrak{y} = \mathfrak{U}\mathfrak{x}$ wieder auf E, da wegen $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$

$$(37)\qquad |\mathfrak{y}|^2 = \sum_{i=1}^{n} |y_i|^2 = \sum_{i=1}^{n} \Big|\sum_{k=1}^{n} u_{ik} x_k\Big|^2 = \sum_{i=1}^{n}\sum_{k=1}^{n} u_{ik} x_k \sum_{j=1}^{n} \overline{u}_{ij} \overline{x}_j$$

$$= \sum_{k=1}^{n}\sum_{j=1}^{n} x_k \overline{x}_j \sum_{i=1}^{n} u_{ik}\overline{u}_{ij} = \sum_{k=1}^{n}\sum_{j=1}^{n} x_k \overline{x}_j \delta_{kj} = \sum_{k=1}^{n} x_k \overline{x}_k = |\mathfrak{x}|^2 = 1$$

gilt. Und zwar ist jeder Punkt von E ein Bildpunkt $\mathfrak{y} = \mathfrak{U}\mathfrak{x}$, da ja $\mathfrak{y} = \mathfrak{U}^{-1}\mathfrak{x}$ gilt und da auch $\mathfrak{U}^{-1}$ unitär ist. Wegen der eindeutigen Umkehrbarkeit entsprechen schließlich verschiedenen Punkten $\mathfrak{x}$ verschiedene Bildpunkte $\mathfrak{y}$. Zusammenfassend können wir sagen, daß durch eine unitäre Transformation von $\mathfrak{x}$ die Mannigfaltigkeit E eineindeutig in sich übergeführt wird. Übrigens ist die Bedingung $|\mathfrak{U}\mathfrak{x}|^2 \equiv |\mathfrak{x}|^2$ wegen (37) nicht nur notwendig, sondern auch hinreichend dafür, daß $\mathfrak{U}$ unitär ist.

Der Übergang von $\mathfrak{x}$ zu $\mathfrak{U}\mathfrak{x}$ bedeutet für $\mathfrak{A}$ nach § 7 den Übergang von $\mathfrak{A}$ zu $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$. Da bei dem Übergang die Mannigfaltigkeit E eineindeutig in sich übergeht, so ist der Wertevorrat von $\mathfrak{A}$ unitären Transformationen gegenüber invariant, also $\mathsf{W}(\mathfrak{A}) = \mathsf{W}(\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1})$. Freilich kann (bereits aus Dimensionsgründen) die Kopplungsform $\Phi(\mathfrak{A})$ denselben Wert auch in verschiedenen Punkten von E annehmen.

§ 18. Der Hausdorffsche Satz über die Konvexität des Wertbereiches.

Toeplitz hat bewiesen, daß die äußere Berandung der Punktmenge $\mathsf{W}(\mathfrak{A})$ der Φ-Ebene konvex ist, und er hat vermutet, daß $\mathsf{W}(\mathfrak{A})$ nicht hohl sein kann, daß also $\mathsf{W}(\mathfrak{A})$ ein konvexer Bereich ist. [Eine abgeschlossene und beschränkte ebene Punktmenge heißt bekanntlich ein konvexer

Bereich, wenn sie von einer Geraden der Ebene entweder überhaupt nicht oder in einem einzigen Punkte getroffen, oder aber entlang eines zusammenhängenden Intervalles, das mit dem Bereich auch identisch sein kann, geschnitten wird.] Hausdorff hat diese Toeplitzsche Vermutung auf die folgende Weise bewiesen.

Man setze $\mathfrak{A} = \mathfrak{A}_1 + \sqrt{-1}\,\mathfrak{A}_2$, wobei $\mathfrak{A}_1$, $\mathfrak{A}_2$ die eindeutig bestimmten Hermiteschen Komponenten von $\mathfrak{A}$ bezeichnen (siehe § 17). Dann unterwerfe man $\mathfrak{A}$ einer unitären Transformation $\mathfrak{U}$, die so beschaffen ist, daß $\mathfrak{U}\mathfrak{A}_1\mathfrak{U}^{-1}$ eine Diagonalmatrix wird. Man hat $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1} = \mathfrak{U}\mathfrak{A}_1\mathfrak{U}^{-1} + \sqrt{-1}\,\mathfrak{U}\mathfrak{A}_2\mathfrak{U}^{-1}$, wobei $\mathfrak{U}\mathfrak{A}_1\mathfrak{U}^{-1}$, $\mathfrak{U}\mathfrak{A}_2\mathfrak{U}^{-1}$ nach § 15 Hermitesch sind, also die eindeutig bestimmten Hermiteschen Komponenten von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ darstellen. Nun haben aber $\mathfrak{A}$ und $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ nach § 17 denselben Wertevorrat. Wir können also annehmen, daß die unitäre Transformation, durch welche $\mathfrak{A}_1$ auf die Hauptachsen gedreht wird, bereits ausgeübt ist, d. h. wir können $\mathfrak{U} = \mathfrak{E}$ voraussetzen.

Offenbar gilt

$$\Phi(\mathfrak{A}; \mathfrak{x}) = \Phi(\mathfrak{A}_1; \mathfrak{x}) + \sqrt{-1}\,\Phi(\mathfrak{A}_2; \mathfrak{x}). \tag{38}$$

Ferner ist der Wertevorrat einer Hermiteschen Matrix, wie leicht eingesehen werden kann, auf der reellen Achse der Φ-Ebene gelegen. Denn wegen $a_{ik} = \bar{a}_{ki}$ ist die Zahl $a_{ik} x_i \bar{x}_k + a_{ki} x_k \bar{x}_i$ zu sich selbst konjugiert, also ist sie reell, mithin muß offenbar auch $\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik} x_i \bar{x}_k$ reell sein. Und zwar ist der Wertevorrat entweder ein einziger Punkt oder ein zusammenhängendes, beiderseits abgeschlossenes, endliches Intervall. Denn er ist ein stetiges Bild der Mannigfaltigkeit $\sum_{i=1}^{n} |x_i|^2 = 1$ (vgl. weiter unten). Setzen wir

$$\begin{gathered} \Phi = \Phi_1 + \sqrt{-1}\,\Phi_2; \\ \Phi = \Phi(\mathfrak{A}; \mathfrak{x}), \quad \Phi_1 = \Phi(\mathfrak{A}_1; \mathfrak{x}), \quad \Phi_2 = \Phi(\mathfrak{A}_2; \mathfrak{x}), \end{gathered} \tag{39}$$

so ist, da $\mathfrak{A}_1$ und $\mathfrak{A}_2$ Hermitesch sind, Φ_1 die reelle, Φ_2 die imaginäre Komponente der Zahl Φ. Es bezeichne $\mathrm{\Phi}_1$ irgendeine festgewählte, im Wertevorrat von $\mathfrak{A}_1$ gelegene Zahl, d. h. eine Zahl, die die reelle Komponente einer im Wertevorrat von $\mathfrak{A}$ gelegenen Zahl ist. Es bezeichne ferner $\Gamma(\mathrm{\Phi}_1)$ diejenige Teilmenge der Mannigfaltigkeit E, auf welcher $\Phi(\mathfrak{A}_1; \mathfrak{x}) = \mathrm{\Phi}_1$ wird. Es seien

$$\mathfrak{y} = (\eta_1, \eta_2, \ldots, \eta_n), \qquad \mathfrak{z} = (\zeta_1, \zeta_2, \ldots, \zeta_n)$$

irgend zwei Punkte von $\Gamma(\mathrm{\Phi}_1)$. Wir wollen zeigen, daß diese beiden Punkte mit einem auf $\Gamma(\mathrm{\Phi}_1)$ verlaufenden, stetigen (komplexen) Kurvenzug verbunden werden können, daß also die Mannigfaltigkeit $\Gamma(\mathrm{\Phi}_1)$

zusammenhängend ist. Wir können dabei voraussetzen, daß $\mathfrak{y} \neq \mathfrak{z}$ gilt. Man setze $\eta_i = |\eta_i| e^{\varphi_i}$, $\zeta_i = |\zeta_i| e^{\psi_i}$, so daß φ_i den mit $\sqrt{-1}$ multiplizierten Bogen von η_i darstellt, also, ebenso wie ψ_i, Null oder rein imaginär ist. Läuft der reelle Parameter t von 0 bis 1, so werden durch die drei stetigen Kurven

$$(K_1) \qquad x_i^{(1)}(t) = |\eta_i| e^{(1-t)\varphi_i}, \qquad i = 1, 2, \ldots, n;$$

$$(K_2) \qquad x_j^{(2)}(t) = \sqrt{(1-t)|\eta_i|^2 + t|\zeta_i|^2}, \qquad i = 1, 2, \ldots, n;$$

$$(K_3) \qquad x_i^{(3)}(t) = |\zeta_i| e^{t\psi_i}, \qquad i = 1, 2, \ldots, n$$

bzw. die drei Punktepaare

$$(P_1) \qquad \mathfrak{y} = (\eta_1, \eta_2, \ldots, \eta_n), \ (|\eta_1|, |\eta_2|, \ldots, |\eta_n|);$$

$$(P_2) \qquad (|\eta_1|, |\eta_2|, \ldots, |\eta_n|), \ (|\zeta_1|, |\zeta_2|, \ldots, |\zeta_n|);$$

$$(P_3) \qquad (|\zeta_1|, |\zeta_2|, \ldots, |\zeta_n|), \ (\zeta_1, \zeta_2, \ldots, \zeta_n) = \mathfrak{z}$$

gewiß verbunden, da ja $x_i^{(1)}(0) = \eta_i$, $x_i^{(1)}(1) = |\eta_i|$ usf. gilt. Und zwar endet K_1 in dem Punkte, in dem K_2 anfängt, und K_3 geht von dem Endpunkte von K_2 aus, während K_1 von $\mathfrak{y}$ ausgeht und K_3 in $\mathfrak{z}$ mündet, so daß durch den stetigen Kurvenzug $K_1 + K_2 + K_3$ die beiden Punkte $\mathfrak{y}, \mathfrak{z}$ von E gewiß verbunden werden. Wegen $\sum_{i=1}^{n} |\mathfrak{y}_i|^2 = \sum_{i=1}^{n} |\zeta_i|^2 = 1$ leuchtet ferner unmittelbar ein, daß $\sum_{i=1}^{n} |x_i^{(\nu)}(t)|^2 \equiv 1$ für alle drei ν gilt, so daß $K_1 + K_2 + K_3$ auf E gelegen ist. Darüber hinaus haben wir zu zeigen, daß $K_1 + K_2 + K_3$ sogar in der Teilmenge $\Gamma(\Phi_1)$ verläuft. Nun ist $\mathfrak{A}_1$ nach Voraussetzung eine Diagonalmatrix, etwa $= \|\alpha_i \delta_{ik}\|$, also ist ihre Kopplungsform

$$(40) \qquad \Phi(\mathfrak{A}_1; \mathfrak{x}) = \Phi(\mathfrak{A}_1; x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \alpha_i x_i \bar{x}_i = \sum_{i=1}^{n} \alpha_i |x_i|^2.$$

Da die Punkte $\mathfrak{y}$ und $\mathfrak{z}$ nach Voraussetzung auf $\Gamma(\Phi_1)$ liegen, so daß $\Phi(\mathfrak{A}_1; \mathfrak{y}) = \Phi(\mathfrak{A}_1; \mathfrak{z}) = \Phi_1$ gilt, so folgt daher mit Rücksicht auf die Gleichungen der drei Kurven K, daß für alle t

$$\Phi(\mathfrak{A}_1; \mathfrak{x}^{(1)}(t)) = \sum_{i=1}^{n} |x_i^{(1)}(t)|^2 \alpha_i = \sum_{i=1}^{n} \alpha_i |\eta_i|^2 = \Phi(\mathfrak{A}_1; \mathfrak{y}) = \Phi_1,$$

$$\Phi(\mathfrak{A}_1; \mathfrak{x}^{(2)}(t)) = \sum_{i=1}^{n} \alpha_i |x_i^{(2)}(t)|^2 = t \sum_{i=1}^{n} \alpha_i |\eta_i|^2 + (1-t) \sum_{i=1}^{n} \alpha_i |\zeta_i|^2$$

$$= t\,\Phi(\mathfrak{A}_1; \mathfrak{y}) + (1-t)\,\Phi(\mathfrak{A}_1; \mathfrak{z}) = t\,\Phi_1 + (1-t)\,\Phi_1 = \Phi_1,$$

$$\Phi(\mathfrak{A}_1; \mathfrak{x}^{(3)}(t)) = \sum_{i=1}^{n} \alpha_i |x_i^{(3)}(t)|^2 = \sum_{i=1}^{n} \alpha_i |\zeta_i|^2 = \Phi(\mathfrak{A}_1; \mathfrak{z}) = \Phi_1$$

gilt, so daß die Kopplungsform von $\mathfrak{A}_1$ längs $K_1 + K_2 + K_3$ denselben Wert Φ_1 wie in den beiden Endpunkten annimmt und daher $K_1 + K_2 + K_3$ nicht nur auf E, sondern auch auf $\Gamma(\Phi_1)$ gelegen ist. Damit ist gezeigt, daß die Teilmenge $\Gamma(\Phi_1)$ von E zusammenhängend ist. Sie ist ferner abgeschlossen, da die Kopplungsform $\Phi(\mathfrak{A}_1; \mathfrak{x})$ stetig und E abgeschlossen ist. Endlich ist E, also auch $\Gamma(\Phi_1)$ beschränkt. Da die Hermitesche Kopplungsform $\Phi(\mathfrak{A}_2; \mathfrak{x})$ nur reeller Werte fähig ist, so folgt daraus nach bekannten Sätzen (von Bolzano-Weierstraß bzw. Heine-Borel), daß sie auf $\Gamma(\Phi_1)$ ein Maximum und ein Minimum hat und alle dazwischen liegenden Werte bereits auf $\Gamma(\Phi_1)$ annimmt (freilich kann das Maximum mit dem Minimum zusammenfallen). Wegen (38), (39) besagt dies (siehe S. 35), daß die in der Φ-Ebene gelegene Punktmenge $\mathsf{W}(\mathfrak{A})$ von jeder der imaginären Achse parallelen Geraden $\Phi = \Phi_1$ entweder überhaupt nicht oder in einem einzigen Punkte getroffen oder längs eines zusammenhängenden Intervalles geschnitten wird. Nun kommt es hier auf die Richtung der Geraden überhaupt nicht an. Denn bezeichnet γ eine reelle Zahl, so entsteht die Wertevorratspunktmenge $\mathsf{W}(e^{\gamma\sqrt{-1}}\mathfrak{A})$ der Matrix $e^{\gamma\sqrt{-1}}\mathfrak{A}$ aus der zu $\mathfrak{A}$ gehörigen Punktmenge $\mathsf{W}(\mathfrak{A})$ offenbar durch eine Drehung mit dem Winkel γ um den Punkt $\Phi = 0$. Damit ist der Hausdorffsche Satz bewiesen.

§ 19. Wertbereich und Spektrum. Die Toeplitzsche Polygonregel.

Durch den Wertevorrat von $\mathfrak{A}$ werden die Eigenwerte von $\mathfrak{A}$ „von oben“ eingeschränkt. Z. B. ist jeder Eigenwert dem Betrage nach $\leqq \mathsf{M}(\mathfrak{A})$, die Beträge der reellen bzw. der imaginären Komponenten der Eigenwerte von $\mathfrak{A}$ sind $\leqq \mathsf{M}(\mathfrak{A}_1)$ bzw. $\mathsf{M}(\mathfrak{A}_2)$, wenn $\mathfrak{A}_1$ und $\mathfrak{A}_2$ die Hermiteschen Komponenten von $\mathfrak{A}$ bezeichnen. Genauer läßt sich nach Toeplitz folgendes behaupten: deutet man die n Eigenwerte von $\mathfrak{A}$ ebenfalls in der Φ-Ebene, so liegen sie alle in dem Bereiche $\mathsf{W}(\mathfrak{A})$ des Wertevorrates (möglicherweise auf seinem Rande). Ist nämlich λ ein Eigenwert, so gibt es einen Vektor $\mathfrak{x}$ mit $\mathfrak{x}\bar{\mathfrak{x}} = 1$ (d. h. auf E) derart, daß $(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{o}$, also $\lambda\mathfrak{x} = \mathfrak{A}\mathfrak{x}$, mithin $\lambda\mathfrak{x}\bar{\mathfrak{x}} = {}^{-}\mathfrak{A}\mathfrak{x} = \Phi(\mathfrak{A}; \mathfrak{x})$, d. h $\lambda = \Phi(\mathfrak{A}; \mathfrak{x})$ gilt, wobei $\mathfrak{x}$ auf E, also $\Phi(\mathfrak{A}; \mathfrak{x})$ und daher auch λ auf $\mathsf{W}(\mathfrak{A})$ liegt, w. z. b. w.

Die naheliegende umgekehrte Frage, inwiefern durch die Konfiguration der Eigenwerte diejenige des Wertevorrates eingeschränkt wird, hat bei *nicht normalen* Matrizen keinen Sinn. Man bedenke nur, daß eine nicht normale Matrix durch die Eigenwerte nicht bis

auf eine unitäre Transformation festgelegt wird, und daß dabei das, was in der kanonischen Darstellung unterhalb der Diagonale liegt, völlig willkürlich bleibt. Man denke an den Fall, wo sogar alle Eigenwerte verschwinden. — Für die *normalen* Matrizen wird der Zusammenhang des Wertevorrates mit den Eigenwerten durch den folgenden Toeplitzschen Satz restlos aufgedeckt: Der Wertevorrat $\mathsf{W}(\mathfrak{A})$ einer normalen Matrix ist identisch mit dem kleinsten konvexen Bereich $\mathsf{T}(\mathfrak{A})$ der Φ-Ebene, der alle n Eigenwerte von $\mathfrak{A}$ enthält. — Da die Eigenwerte immer in $\mathsf{W}(\mathfrak{A})$ liegen und da $\mathsf{W}(\mathfrak{A})$ konvex ist, so liegt freilich $\mathsf{T}(\mathfrak{A})$ immer in $\mathsf{W}(\mathfrak{A})$, doch wird $\mathsf{W}(\mathfrak{A})$ bei ganz willkürlichen Matrizen über $\mathsf{T}(\mathfrak{A})$ hinausragen, während im normalen Fall, wie sogleich bewiesen werden soll, $\mathsf{W}(\mathfrak{A}) = \mathsf{T}(\mathfrak{A})$ ist. Es ist anschaulich einleuchtend, daß $\mathsf{T}(\mathfrak{A})$ immer ein konvexes Polygon (evtl. ein Intervall oder ein einziger Punkt) ist, dessen sämtliche Eckpunkte gewiß Eigenwerte von $\mathfrak{A}$ sind. Man kann $\mathsf{T}(\mathfrak{A})$ auf bekannte Weise mit einem umschlingenden Faden „konstruieren".

Es sei $\mathfrak{A}$ normal. Da $\mathsf{W}(\mathfrak{A})$ unitär invariant ist, können wir annehmen, daß $\mathfrak{A}$ bereits auf die Diagonalform gebracht wurde, so daß $\mathfrak{A} = \|\lambda_i \delta_{ik}\|$, also die Kopplungsform $= \lambda_1 x_1 \bar{x}_1 + \lambda_2 x_2 \bar{x}_2 + \cdots$, d. h. $\Phi(\mathfrak{A}; \mathfrak{y}) = \sum_{i=1}^{n} \lambda_i |x_i|^2$ ist. Hierbei ist $\sum_{i=1}^{n} |x_i|^2 = 1$. Folglich sind die beiden Komponenten der komplexen Zahl $\Phi(\mathfrak{A}; \mathfrak{x})$ die Koordinaten des Schwerpunktes einer Gewichtsverteilung in der Φ-Ebene: der Punkt $\Phi = \lambda_i$ ist mit dem Gewicht $|x_i|^2 \geqq 0$ derart belastet, daß das Gesamtgewicht die Einheit beträgt (die λ_i brauchen nicht verschieden zu sein). Es leuchtet anschaulich ein und genauer folgt es nach bekannten, auf Weierstraß und Minkowski zurückgehenden Betrachtungen[1]), daß einerseits der Schwerpunkt in $\mathsf{T}(\mathfrak{A})$ liegt, und daß andererseits, bei passender Verteilung des Gesamtgewichtes Eins auf die Punkte λ_i, jeder Punkt von $\mathsf{T}(\mathfrak{A})$ zum Schwerpunkt werden kann[2]). Da der Schwerpunkt durch $\Phi(\mathfrak{A}; \mathfrak{y})$ dargestellt wird, so besagt dies, daß $\mathsf{T}(\mathfrak{A}) = \mathsf{W}(\mathfrak{A})$ ist, w. z. b. w.

---

[1]) Vgl. insb. C. Carathéodory, Rend. Palermo **32** (1911), S. 195ff.

[2]) Diesen zweiten (schwierigeren) Teil dieses Satzes benötigen wir zum Beweise der Toeplitzschen Polygonregel eigentlich nicht. Denn wir wissen von vornherein, daß das Spektrum (sogar bei nicht normalen Matrizen) in dem Wertbereiche der Kopplungsform liegt (siehe den Anfang dieses Paragraphen).

§ 20. Die Bilinearform.

Es bezeichne F die Mannigfaltigkeit der normierten Vektoren*paare* $\mathfrak{x}, \mathfrak{y}$, d. i. die Gesamtheit der der Bedingung

$$(41) \qquad \mathsf{F}: \quad \sum_{i=1}^{n} |x_i|^2 = 1, \qquad \sum_{i=1}^{n} |y_i|^2 = 1$$

genügenden $n + n$ voneinander unabhängigen reellen oder komplexen Zahlen. Man ordne jeder Matrix $\mathfrak{A}$ die folgende Funktion von $n + n$ Veränderlichen als „Bilinearform von $\mathfrak{A}$" zu:

$$(42) \qquad \Psi = \Psi(\mathfrak{A}) = \Psi(\mathfrak{A}; \mathfrak{x}, \mathfrak{y}) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} x_i y_k .$$

Denkt man die Werte Ψ, derer die Bilinearform von $\mathfrak{A}$ unter der Nebenbedingung F fähig ist, in derselben Φ-Ebene wie die Kopplungsform und bezeichnet man durch $\mathsf{V}(\mathfrak{A})$ das durch die Abbildung

$$(43) \qquad \Phi = \Psi(\mathfrak{A}; \mathfrak{x}, \mathfrak{y})$$

in die Φ-Ebene entworfene Bild von F, so ist in $\mathsf{V}(\mathfrak{A})$ gewiß jeder Punkt von $\mathsf{W}(\mathfrak{A})$ enthalten. Denn die Kopplungsform entsteht aus der Bilinearform durch die Kopplung $\mathfrak{y} = \bar{\mathfrak{x}}$, indem

$$(44) \qquad \Phi(\mathfrak{A}; \mathfrak{x}) \equiv \Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{x}})$$

ist. Enthält $\mathsf{V}(\mathfrak{A})$ den Punkt Φ_0, so ist die ganze um den Nullpunkt mit dem Radius $|\Phi_0|$ geschlagene Kreislinie aus Punkten von $\mathsf{V}(\mathfrak{A})$ gebildet. Um dies einzusehen, genügt es, $\mathfrak{y}$ durch $e^{\gamma\sqrt{-1}}\,\mathfrak{y}$ zu ersetzen, wobei γ eine reelle Zahl, also $\left|e^{\gamma\sqrt{-1}}\,\mathfrak{y}\right|^2 = |\mathfrak{y}|^2 = 1$ ist; in der Tat gilt $\Psi(\mathfrak{A}; \mathfrak{x}, C\,\mathfrak{y}) = C\,\Psi(\mathfrak{A}; \mathfrak{x}, \mathfrak{y})$. Da ferner F zusammenhängend und abgeschlossen und die Funktion $|\Psi(\mathfrak{A}; \mathfrak{x}, \mathfrak{y})| \geqq 0$ stetig ist, so bilden der in dem Wertbereiche der Bilinearform liegenden, um den Nullpunkt geschlagenen Kreislinien ein Kontinuum. M. a. W., es gehören zugleich mit den beiden reellen, nicht negativen Zahlen $\Phi^{(1)}$, $\Phi^{(2)}$ auch alle Punkte der Verbindungsstrecke $\Phi^{(1)} \leqq \Phi \leqq \Phi^{(2)}$ der Menge $\mathsf{V}(\mathfrak{A})$ an. Endlich erkennt man leicht, daß es immer $n + n$ der Bedingung $\sum |x_i|^2 = \sum |y_i|^2 = 1$ Genüge leistende Zahlen x_i, y_i angetroffen werden können derart, daß $\sum\sum a_{ik} x_i y_k = 0$ ausfällt, so daß der Punkt $\Phi = 0$ gewiß $\mathsf{V}(\mathfrak{A})$ angehört. Damit ist gezeigt, daß $\mathsf{V}(\mathfrak{A})$ ein um den Punkt $\Phi = 0$ mit dem Radius $\mathsf{P}(\mathfrak{A})$ geschlagener Kreisbereich ist, wobei $\mathsf{P}(\mathfrak{A})$ die obere Grenze des Betrages von $\Psi(\mathfrak{A}; \mathfrak{x}, \mathfrak{y})$ unter der Nebenbedingung F bezeichnet, also nur für die Nullmatrix $\mathfrak{A} = \mathfrak{O}$ verschwindet.

Da $|\Phi| \leqq \mathsf{M}(\mathfrak{A})$ der kleinste, $\mathsf{W}(\mathfrak{A})$ enthaltende Kreis ist (siehe S. 34) und da $\mathsf{W}(\mathfrak{A})$ in $\mathsf{V}(\mathfrak{A})$ liegt, so hat man $\mathsf{M}(\mathfrak{A}) \leqq \mathsf{P}(\mathfrak{A})$. Merkwürdigerweise kann man umgekehrt $\mathsf{P}(\mathfrak{A})$ mittels $\mathsf{M}(\mathfrak{A})$ abschätzen. Wir zeigen jetzt nämlich nach Hellinger und Toeplitz, daß $\mathsf{P}(\mathfrak{A}) \leqq 2\,\mathsf{M}(\mathfrak{A})$ ist. Es sei zunächst $\mathfrak{A}$ Hermitesch. Es gilt identisch

$$\begin{aligned}\Phi(\mathfrak{A}; \mathfrak{x} \pm \mathfrak{y}) &= \Psi(\mathfrak{A}; \mathfrak{x} \pm \mathfrak{y}, \bar{\mathfrak{x}} \pm \bar{\mathfrak{y}})\\ &= \Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{x}}) \pm \Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{y}}) \pm \Psi(\mathfrak{A}; \bar{\mathfrak{x}}, \mathfrak{y}) + \Psi(\mathfrak{A}; \mathfrak{y}, \bar{\mathfrak{y}})\\ &= \Phi(\mathfrak{A}; \mathfrak{x}) \pm \Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{y}}) \pm \overline{\Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{y}})} + \Phi(\mathfrak{A}; \mathfrak{y}),\end{aligned}$$

mithin

$$\begin{aligned}(45)\quad \Phi(\mathfrak{A}; \mathfrak{x} + \mathfrak{y}) - \Phi(\mathfrak{A}, \mathfrak{x} - \mathfrak{y}) &= 2\,\Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{y}}) + 2\,\overline{\Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{y}})}\\ &= 4\,\mathfrak{R}\,\Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{y}}),\end{aligned}$$

wobei $\mathfrak{R}\,a$ den Realteil der Zahl a bezeichnet. Wir behaupten, daß bei Hermiteschen Matrizen $\mathsf{M}(\mathfrak{A}) = \mathsf{P}(\mathfrak{A})$, oder, was wegen $\mathsf{M}(\mathfrak{A}) \leqq \mathsf{P}(\mathfrak{A})$ dasselbe ist, $\mathsf{P}(\mathfrak{A}) \leqq \mathsf{M}(\mathfrak{A})$ gilt. Nun ist aber $\mathsf{P}(\mathfrak{A})$ die obere Grenze von $|\Psi(\mathfrak{A}; \mathfrak{x}, \mathfrak{y})|$ auf F, oder, was dasselbe ist, die obere Grenze von $|\Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{y}})|$ auf F, oder endlich die obere Grenze des Betrages des Realteiles von $\Psi(\mathfrak{A}; \mathfrak{x}, \bar{\mathfrak{y}})$ auf F, da ja die Punktmenge $\mathsf{V}(\mathfrak{A})$ ein Kreis ist. Unsere Behauptung wird also, mit Rücksicht auf (45), bewiesen sein, wenn wir zeigen, daß auf F

$$(46)\qquad |\Phi(\mathfrak{A}; \mathfrak{x} + \mathfrak{y}) - \Phi(\mathfrak{A}; \mathfrak{x} - \mathfrak{y})| \leqq 4\,\mathsf{M}(\mathfrak{A})$$

gilt. Nun ist aber $\mathsf{M}(\mathfrak{A})$ die obere Grenze von $|\Phi(\mathfrak{A}; \mathfrak{x})|$ auf E, also gilt aus Gründen der Homogenität auch außerhalb E die Ungleichheit $|\Phi(\mathfrak{A}; \mathfrak{x})| \leqq \mathsf{M}(\mathfrak{A})\,|\mathfrak{x}|^2$, so daß der Ausdruck linkerhand in (46) gewiß

$$\begin{aligned}&\leqq \mathsf{M}(\mathfrak{A})\,[|\mathfrak{x} + \mathfrak{y}|^2 + |\mathfrak{x} - \mathfrak{y}|^2]\\ &= \mathsf{M}(\mathfrak{A})\,[(\mathfrak{x} + \mathfrak{y})(\bar{\mathfrak{x}} + \bar{\mathfrak{y}}) + (\mathfrak{x} - \mathfrak{y})(\bar{\mathfrak{x}} - \bar{\mathfrak{y}})]\\ &= \mathsf{M}(\mathfrak{A})\,[\mathfrak{x}\bar{\mathfrak{x}} + \mathfrak{y}\bar{\mathfrak{y}} + \mathfrak{x}\bar{\mathfrak{x}} + \mathfrak{y}\bar{\mathfrak{y}}],\end{aligned}$$

und daher auf F, wie behauptet, $\leqq 4\,\mathsf{M}(\mathfrak{A})$ ausfällt. Für Hermitesche Matrizen gilt also $\mathsf{M}(\mathfrak{A}) = \mathsf{P}(\mathfrak{A})$. — Setzt man die Lösbarkeit des Hauptachsenproblems als bekannt voraus, so kann die Richtigkeit dieser Gleichung ohne Rechnung eingesehen werden, und zwar auch für den Fall, wo $\mathfrak{A}$ nicht Hermitesch, sondern nur normal, d. h. mit einer Diagonal-Matrix unitär äquivalent ist. Denn für diagonale $\mathfrak{A}$ ist $\mathsf{M}(\mathfrak{A}) = \mathsf{P}(\mathfrak{A})$ trivial, und aus den Schlußbemerkungen von § 17 folgt leicht, daß die beiden Zahlen $\mathsf{M}(\mathfrak{A})$, $\mathsf{P}(\mathfrak{A})$ unitär invariant sind:

$$\mathsf{M}(\mathfrak{A}) = \mathsf{M}(\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}), \qquad \mathsf{P}(\mathfrak{A}) = \mathsf{P}(\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}),$$

— Ist nun die Matrix $\mathfrak{A}$ nicht normal, so zerlege man sie in ihre beiden Hermiteschen Komponenten, $\mathfrak{A} = \mathfrak{A}_1 + \sqrt{-1}\,\mathfrak{A}_2$. Offenbar ist $\mathsf{P}(\mathfrak{A}) \leqq \mathsf{P}(\mathfrak{A}_1) + \mathsf{P}(\mathfrak{A}_2)$, also, da $\mathfrak{A}_1$ und $\mathfrak{A}_2$ Hermitesch sind, $\mathsf{P}(\mathfrak{A}) \leqq \mathsf{M}(\mathfrak{A}_1) + \mathsf{M}(\mathfrak{A}_2)$. Nun ist aber $\Phi(\mathfrak{A}_1; \mathfrak{x})$, wie bereits auf S. 35 erwähnt wurde, einfach die reelle Komponente der komplexen Zahl $\Phi(\mathfrak{A}; \mathfrak{x})$, folglich ist $|\Phi(\mathfrak{A}_1; \mathfrak{x})| \leqq |\Phi(\mathfrak{A}; \mathfrak{x})|$, also auch $\mathsf{M}(\mathfrak{A}_1) \leqq \mathsf{M}(\mathfrak{A})$, und ebenso $\mathsf{M}(\mathfrak{A}_2) \leqq \mathsf{M}(\mathfrak{A})$, mithin $\mathsf{P}(\mathfrak{A}) \leqq \mathsf{M}(\mathfrak{A}) + \mathsf{M}(\mathfrak{A})$. Damit ist gezeigt, daß für jede Matrix $\mathfrak{A}$

$$0 \leqq \mathsf{M}(\mathfrak{A}) \leqq \mathsf{P}(\mathfrak{A}) \leqq 2\,\mathsf{M}(\mathfrak{A}) \tag{47}$$

gilt. Mehr läßt sich nicht behaupten. Denn für normale Matrizen ist $\mathsf{M}(\mathfrak{A}) = \mathsf{P}(\mathfrak{A})$, während das Toeplitzsche Beispiel

$$\mathfrak{A} = \begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix} \tag{48}$$

zeigt, daß $\mathsf{P}(\mathfrak{A})$ auch $= 2\,\mathsf{M}(\mathfrak{A})$ sein kann.

§ 21. Definite Matrizen. Das Spektrum der beiden Normmatrizen.

Bevor wir weitergehen, wollen wir einiges über die Hermiteschen Matrizen zusammenstellen. — Da die Hermiteschen Matrizen nur reelle Eigenwerte haben, ist ihr Polygon $\mathsf{T}(\mathfrak{A})$ ein zusammenhängendes Intervall auf der reellen Achse der Φ-Ebene, das von zwei Eigenwerten begrenzt wird, alle Eigenwerte enthält und in dem Falle $\mathfrak{A} = c\,\mathfrak{E}$ und offenbar auch nur in diesem Falle auf einen Punkt, nämlich auf $\Phi = c$, zusammenschrumpft. Da ferner die Hermiteschen Matrizen normal sind, so ist $\mathsf{T}(\mathfrak{A}) = \mathsf{W}(\mathfrak{A})$, so daß also der Wertevorrat ein reelles Intervall (eventuell ein reeller Punkt) ist, das alle Eigenwerte enthält und durch den größten und den kleinsten Eigenwert begrenzt wird. Bezeichnet also $\mathsf{m}(\mathfrak{A})$ den größten, $\mathsf{n}(\mathfrak{A})$ den kleinsten Eigenwert von $\mathfrak{A}$ (es können beide auch negativ sein), so ist $\mathsf{m}(\mathfrak{A})$ die obere, $\mathsf{n}(\mathfrak{A})$ die untere Grenze des Wertevorrates von $\mathfrak{A}$; d. h. auf E gilt

$$\mathsf{n}(\mathfrak{A}) \leqq \Phi(\mathfrak{A}; \mathfrak{x}) \leqq \mathsf{m}(\mathfrak{A}), \tag{49}$$

wobei die Schranken durch keine besseren ersetzt werden können, und man hat

$$\mathsf{P}(\mathfrak{A}) = \mathsf{M}(\mathfrak{A}) = \max\left(|\mathsf{m}(\mathfrak{A})|, |\mathsf{n}(\mathfrak{A})|\right), \tag{50}$$

unter $\max(\alpha, \beta) = \max(\beta, \alpha)$ die Zahl β verstanden, wenn $\beta \geqq \alpha$ ist. Es sei betont, daß die Zahlen $\mathsf{m}(\mathfrak{A})$, $\mathsf{n}(\mathfrak{A})$ nur für Hermitesche Matrizen $\mathfrak{A}$ erklärt sind.

Wir zeigen nun, daß bei *jeder* Matrix $\mathfrak{A}$ die (eventuell verschiedenen) Normen $\mathfrak{A}^*\mathfrak{A}$, $\mathfrak{A}\mathfrak{A}^*$ nichtnegativ definite Hermitesche Matrizen sind. Da auch $\mathfrak{A}^*$ ein $\mathfrak{A}$ ist, können wir uns auf die vordere Norm $\mathfrak{A}^*\mathfrak{A}$ beschränken. Sie ist, wie wir wissen (S. 29), immer Hermitesch. Ferner ist ihre Kopplungsform

$$
\begin{aligned}
\Phi(\mathfrak{A}^*\mathfrak{A}; \mathfrak{x}) &= \sum_{i=1}^{n}\sum_{k=1}^{n} x_i \bar{x}_k \sum_{j=1}^{n} \bar{a}_{ji} a_{jk} \\
&= \sum_{j=1}^{n}\Big(\sum_{i=1}^{n} \bar{a}_{ji} x_i\Big)\Big(\sum_{k=1}^{n} a_{jk} \bar{x}_k\Big) = \sum_{j=1}^{n}\Big|\sum_{i=1}^{n} \bar{a}_{ji} x_i\Big|^2 \geqq 0,
\end{aligned}
\tag{51}
$$

also ist $\mathsf{W}(\mathfrak{A}^*\mathfrak{A})$ auf der nichtnegativen reellen Halbachse der Φ-Ebene gelegen. Da $\mathfrak{A}^*\mathfrak{A}$ normal, also $\mathsf{W}(\mathfrak{A}^*\mathfrak{A}) = \mathsf{T}(\mathfrak{A}^*\mathfrak{A})$ ist, so gilt dies auch von den Eigenwerten, die ja den Bereich $\mathsf{T}(\mathfrak{A}^*\mathfrak{A})$ als den kleinsten, sie enthaltenden konvexen Bereich bestimmen. Mithin ist $\mathfrak{A}^*\mathfrak{A}$ eine nichtnegativ definite Hermitesche Matrix. Positiv definit ist die Matrix $\mathfrak{A}^*\mathfrak{A}$ dann und nur dann, wenn alle ihre Eigenwerte > 0, d. h. $\neq 0$ sind, m. a. W. wenn $\mathfrak{A}^*\mathfrak{A}$ nicht entartet ist. Entsprechendes gilt von $\mathfrak{A}\mathfrak{A}^*$.

Ist $\mathfrak{A}$ eine beliebige Matrix, so ist ihre Bilinearform auf F nach der Schwarzschen Ungleichung derart beschaffen, daß

$$
\begin{aligned}
\Big|\sum_{i=1}^{n} x_i \sum_{k=1}^{n} a_{ik} y_k\Big|^2 &\leqq \sum_{i=1}^{n} |x_i|^2 \sum_{i=1}^{n}\Big|\sum_{k=1}^{n} a_{ik} y_k\Big|^2 \\
&= \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik} y_k \sum_{j=1}^{n} \bar{a}_{ij} \bar{y}_j = \sum_{k=1}^{n}\sum_{j=1}^{n} y_k \bar{y}_j \sum_{i=1}^{n} a_{ik} \bar{a}_{ij},
\end{aligned}
$$

also offenbar

$$
|\Psi(\mathfrak{A}; \mathfrak{x}, \mathfrak{y})|^2 \leqq \Phi(\mathfrak{A}^*\mathfrak{A}; \mathfrak{y}) \tag{52}
$$

und daher auch $[\mathsf{P}(\mathfrak{A})]^2 \leqq \mathsf{M}(\mathfrak{A}^*\mathfrak{A})$ ist. Nun muß hierbei das Gleichheitszeichen gelten. Um dies zu zeigen, genügt es nachzuweisen, daß bei jedem festen $\mathfrak{y}$ ein Punkt $\mathfrak{x}$ auf E vorhanden ist derart, daß

$$
\Big|\sum_{i=1}^{n} x_i A_i\Big|^2 = \sum_{i=1}^{n} |x_i|^2 \sum_{i=1}^{n} |A_i|^2 \qquad \Big(A_i = \sum_{k=1}^{n} a_{ik} y_k\Big)
$$

gilt. Dies ist aber (aus Gründen der Homogenität) mit der auf S. 10 erwähnten Umkehrung der Schwarzschen Ungleichheit äquivalent. — Mithin ist

$$
[\mathsf{P}(\mathfrak{A})]^2 = \mathsf{M}(\mathfrak{A}^*\mathfrak{A}). \tag{53}
$$

Wendet man dies auf die Matrix $\mathfrak{A}^*$ an, so folgt wegen $\mathfrak{A}^{**} = \mathfrak{A}$

(54) $$[\mathsf{P}(\mathfrak{A}^*)]^2 = \mathsf{M}(\mathfrak{A}\mathfrak{A}^*).$$

Da $\mathsf{P}(\mathfrak{A})$ die obere Grenze des Betrages der Bilinearform von $\mathfrak{A}$ auf F bezeichnet, so ist offenbar $\mathsf{P}(\mathfrak{A}) = \mathsf{P}(\mathfrak{A}')$, aber auch $= \mathsf{P}(\overline{\mathfrak{A}})$, und daher auch $= \mathsf{P}(\mathfrak{A}^*)$, also nach (53), (54)

(55) $$\mathsf{M}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{M}(\mathfrak{A}^*\mathfrak{A}),$$

wie auch die Matrix $\mathfrak{A}$ sein mag. Ist $\mathfrak{H}$ Hermitesch, so bezeichnet $\mathsf{m}(\mathfrak{H})$ die obere, $\mathsf{n}(\mathfrak{H})$ die untere Grenze von $\Phi(\mathfrak{H})$ auf E, während $\mathsf{M}(\mathfrak{H})$ die obere Grenze des absoluten Betrages von $\Phi(\mathfrak{H})$ unter derselben Nebenbedingung bedeutet. Ist also $\mathfrak{H}$ eine nichtnegativ definite Hermitesche Matrix, so ist $0 \leqq \mathsf{n}(\mathfrak{H}) \leqq \mathsf{m}(\mathfrak{H}) = \mathsf{M}(\mathfrak{H})$. Setzt man hierin $\mathfrak{H} = \mathfrak{A}^*\mathfrak{A}$ und sodann $\mathfrak{H} = \mathfrak{A}\mathfrak{A}^*$, was gestattet ist, da die beiden Normen einer jeden Matrix nichtnegativ definite Hermitesche Matrizen sind, so folgt wegen (55)

(56) $$\mathsf{m}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{m}(\mathfrak{A}^*\mathfrak{A})$$

für jede Matrix $\mathfrak{A}$.

Um eine entsprechende Beziehung auch für die n herzuleiten, bemerken wir zuerst, daß für jede Matrix $\mathfrak{A}$

(57) $$\det(\mathfrak{A}\mathfrak{A}^*) = \det(\mathfrak{A}^*\mathfrak{A}) = |\det\mathfrak{A}|^2$$

gilt. In der Tat ist nach dem Multiplikationssatz der Determinanten $\det(\mathfrak{A}\mathfrak{A}^*) = (\det\mathfrak{A})(\det\mathfrak{A}^*)$ und die beiden Faktoren rechterhand sind wegen

(58) $$\det\mathfrak{A} = \det\mathfrak{A}', \qquad \det\overline{\mathfrak{A}} = \overline{(\det\mathfrak{A})}, \qquad \mathfrak{A}^* = \overline{\mathfrak{A}}'$$

zueinander konjugiert. Aus (57) folgt, daß alle drei Matrizen $\mathfrak{A}$, $\mathfrak{A}^*\mathfrak{A}$, $\mathfrak{A}\mathfrak{A}^*$ dann und nur dann entartet (d. h. von verschwindender Determinante) sind, wenn dies für mindestens eine von ihnen zutrifft. Betrachten wir zunächst den Fall, wo $\mathfrak{A}$ nicht entartet ist. Dann haben auch $\mathfrak{A}\mathfrak{A}^*$ und $\mathfrak{A}^*\mathfrak{A}$ Reziproken. Und zwar sind die nichtnegativ definiten Hermiteschen Matrizen $\mathfrak{A}\mathfrak{A}^*$, $\mathfrak{A}^*\mathfrak{A}$ sogar positiv definit. Denn würde ein Eigenwert verschwinden, so würde die Determinante $=0$ sein und es wäre keine Reziproke vorhanden. Mithin sind $(\mathfrak{A}\mathfrak{A}^*)^{-1}$ und $(\mathfrak{A}^*\mathfrak{A})^{-1}$ nach § 30 positiv definite Hermitesche Matrizen. Die Eigenwerte von $(\mathfrak{A}\mathfrak{A}^*)^{-1}$ sind die reziproken Eigenwerte von $\mathfrak{A}\mathfrak{A}^*$, mithin ist der kleinste Eigenwert von $\mathfrak{A}\mathfrak{A}^*$ gleich dem reziproken

Werte des größten Eigenwertes von $(\mathfrak{A}\mathfrak{A}^*)^{-1}$, also ist

$$\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) = \frac{1}{\mathsf{m}((\mathfrak{A}\mathfrak{A}^*)^{-1})},$$

und ebenso

$$\mathsf{n}(\mathfrak{A}^*\mathfrak{A}) = \frac{1}{\mathsf{m}((\mathfrak{A}^*\mathfrak{A})^{-1})}.$$

Nun ist aber $(\mathfrak{A}^*\mathfrak{A})^{-1} = \mathfrak{A}^{-1}\mathfrak{A}^{*-1} = \mathfrak{A}^{-1}\mathfrak{A}^{-1*}$ und ebenso $(\mathfrak{A}\mathfrak{A}^*)^{-1} = \mathfrak{A}^{-1*}\mathfrak{A}^{-1}$, also

$$\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) = \frac{1}{\mathsf{m}(\mathfrak{A}^{-1*}\mathfrak{A}^{-1})}, \qquad \mathsf{n}(\mathfrak{A}^*\mathfrak{A}) = \frac{1}{\mathsf{m}(\mathfrak{A}^{-1}\mathfrak{A}^{-1*})},$$

wobei die beiden Nenner identisch sind, da (56) für jede Matrix, also auch für $\mathfrak{A}^{-1}$ gilt. Mithin ist

$$\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{n}(\mathfrak{A}^*\mathfrak{A}). \tag{59}$$

Es war bisher vorausgesetzt, daß $\mathfrak{A}$ nicht entartet ist. Ist aber $\mathfrak{A}$ entartet, so sind es auch die beiden Normen, sie haben also beide mindestens einen verschwindenden Eigenwert, und zwar ist die Null der *kleinste* Eigenwert bei den beiden Normen, da ja die Eigenwerte dieser nichtnegativ definiten Hermiteschen Matrizen gewiß $\geqq 0$ sind. Folglich sind die beiden unter (59) stehenden Zahlen $= 0$, so daß (59) auch dann gilt, wenn $\mathfrak{A}$ entartet ist.

Die beiden Gleichungen (56), (59) sagen es aus, daß die beiden Normen einer jeden Matrix denselben größten und auch denselben kleinsten Eigenwert haben[1]). Also haben die beiden Normkopplungsformen $\Phi(\mathfrak{A}\mathfrak{A}^*; \mathfrak{x})$, $\Phi(\mathfrak{A}^*\mathfrak{A}; \mathfrak{x})$, die immer Hermitesch sind, bei jedem $\mathfrak{A}$ auf E denselben Wertevorrat $\mathsf{W}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{W}(\mathfrak{A}^*\mathfrak{A})$, der ein Intervall oder ein Punkt ist. Bei unitären Matrizen, $\mathfrak{A} = \mathfrak{U}$, ist er wegen $\mathfrak{U}\mathfrak{U}^* = \mathfrak{U}^*\mathfrak{U} = \mathfrak{E}$ immer der Punkt $\Phi = 1$, weil die Punktmege $\mathsf{W}(\mathfrak{E})$ mit diesem einen Punkt identisch ist.

Zum Schluß wollen wir hier für spätere Zwecke eine Ungleichung notieren. — Ist die Hermitesche Matrix $\mathfrak{A}$ nichtnegativ definit, so sind ihre Eigenwerte $\geqq 0$, und da sie normal ist, ist ihre Kopplungsform

$$\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}\, x_i \bar{x}_k \geqq 0$$

[1]) Dies, und auch mehr, wissen wir bereits aus § 14 her. Doch könnten wir den sich aus § 14 ergebenden Beweis für (55) und (56) nicht auf unendliche Matrizen übertragen. Denn es wird bei den unendlichen Matrizen gar nicht selbstverständlich sein, daß z. B. (56) die größte, in dem Spektrum von $\mathfrak{A}\mathfrak{A}^*$ oder $\mathfrak{A}^*\mathfrak{A}$ gelegene Zahl ist. Vgl. §§ 64, 65.

(siehe § 17). Setzt man hierin außer x_i alle übrigen Variablen $= 0$ und $x_i = 1$, so bleibt $a_{ii} \geqq 0$. Setzt man außer x_i und x_k $(i \neq k)$ die übrigen gleich Null, so folgt

$$a_{ii} x_i \bar{x}_i + a_{ik} x_i \bar{x}_k + a_{ki} x_k \bar{x}_i + a_{kk} x_k \bar{x}_k \geqq 0.$$

Linkerhand steht die Kopplungsform der binären Hermiteschen Matrix

$$\begin{Vmatrix} a_{ii} & a_{ik} \\ a_{ki} & a_{kk} \end{Vmatrix};$$

da ihr Wertevorrat durchweg $\geqq 0$ ist und da der Wertevorrat alle Eigenwerte enthält, so sind die beiden Eigenwerte der binären Matrix $\geqq 0$, also ist auch die Determinante $a_{ii} a_{kk} - a_{ik} a_{ki}$, als Produkt der beiden Eigenwerte, $\geqq 0$. Es gilt daher für jede nichtnegativ definite Hermitesche Matrix $\| a_{ik} \|$ die Ungleichheit

(60) $$0 \leqq a_{ik} a_{ki} = | a_{ik} |^2 \leqq a_{ii} a_{kk}; \qquad a_{ii} \geqq 0.$$

§ 22. Unitäre Matrizen. Nichtspektrale Grundgebiete.

Wir wollen hier einiges über die unitären Matrizen zusammenstellen. $\mathfrak{U}$ ist unitär, wenn $\mathfrak{U}^{-1}$ vorhanden und $= \mathfrak{U}^*$ ist. Da mit ihrer Reziproken jede (endliche) Matrix vertauschbar ist, so sind die unitären Matrizen gewiß normal. Es war bereits erwähnt, daß das Produkt von unitären Matrizen wieder unitär ist. Ferner ist auch die Reziproke jeder unitären Matrix unitär. Transformiert man daher eine unitäre Matrix unitär, so erhält man eine unitäre Matrix. Um also das Spektrum zu untersuchen, können wir annehmen, daß $\mathfrak{U}$ eine Diagonalmatrix, $= \| \lambda_i \delta_{ik} \|$ ist. Dann folgt $\mathfrak{U}^{-1} = \mathfrak{U}^* = \| \bar{\lambda}_i \delta_{ik} \|$, also $\mathfrak{E} = \mathfrak{U}\mathfrak{U}^{-1} = \| \lambda_i \bar{\lambda}_i \delta_{ik} \|$, mithin $\| \delta_{ik} \| = \| |\lambda_i|^2 \delta_{ik} \|$, d. h. $|\lambda_i|^2 = 1$, so daß alle Eigenwerte der unitären Matrizen vom Betrage Eins sind (insbesondere ist also $|\det \mathfrak{U}| = 1$). Umgekehrt ist jede *normale* Matrix, deren (nicht notwendig verschiedenen) Eigenwerte alle vom Betrage Eins sind, gewiß unitär. Aus $\mathfrak{U} = \| \lambda_i \delta_{ik} \|$ und $|\lambda_i| = 1$ folgt nämlich wegen $\bar{\lambda}_i = \frac{1}{\lambda_i}$ zunächst $\mathfrak{U}^* = \left\| \frac{\delta_{ik}}{\lambda_i} \right\|$, und daraus $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$. — Indem man die unitäre Transformation der unitären Matrizen auf die Diagonalform etwa nach dem Schur-Schmidtschen Verfahren ausführt, wird man zwanglos zu einer von Hurwitz herrührenden Parameterdarstellung aller unitären Matrizen geführt, deren Parameter den z. B. aus der Mechanik bekannten Eulerschen Winkeln reeller Drehungen entsprechen. — Der Wertevorrat $\mathsf{W}(\mathfrak{U})$ ist nach der Toeplitzschen Regel

(§ 19) ein konvexes Polygon mit höchstens n Eckpunkten, die alle auf dem Rande des Einheitskreises liegen. Gibt es nur zwei verschiedene unter den n Eigenwerten, so tritt an Stelle des Polygons eine Sehne, die dann und nur dann auf einen Punkt zusammenschrumpft, wenn $\mathfrak{U} = c\,\mathfrak{E}$, $|c| = 1$ ist. — Um auch $\mathsf{V}(\mathfrak{U})$ zu beschreiben, beachte man nur, daß, wie man sich leicht überzeugt, wieder $\mathfrak{U} = \|\lambda_i \delta_{ik}\|$ angenommen werden kann. Dann folgt wegen $|\lambda_i| = 1$

$$|\Psi(\mathfrak{U}; \mathfrak{x}, \mathfrak{y})| = \Big|\sum_{i=1}^{n} \lambda_i x_i y_i\Big| \leqq \sum_{i=1}^{n} |x_i y_i| \leqq \sum_{i=1}^{n} |x_i|^2 \sum_{i=1}^{n} |y_i|^2 = 1,$$

während die Eigenwerte nach §§ 19, 20 gewiß auf F liegen. Mithin ist der Kreisbereich $\mathsf{V}(\mathfrak{U})$ der Einheitskreis.

Wir haben den Wertevorrat, der zu $\mathfrak{A}$ gehört, unter der Nebenbedingung $\sum_{1}^{n} |x_i|^2 = 1$ betrachtet und dabei gefunden, daß diese „Kugelbedingung" den spektralen Verhältnissen angepaßt ist. Es sind jedoch für andere Zwecke andere Normalgebiete erforderlich, z. B. die Zugrundelegung der oktaederartigen Bildung $\sum_{i=1}^{n} |x_i| = 1$, oder allgemeiner der Bedingung $\sum_{i=1}^{n} |x_i|^{\delta} = 1$ mit $\delta > 0$, wovon die „Würfelbedingung" $|x_1| = |x_2| = \cdots = |x_n| = 1$ gewissermaßen der Grenzfall ist $(\delta \to \infty)$. Die (insbesondere asymptotische) Untersuchung des gegenseitigen Verhältnisses der zugehörigen Wertevorräte von $\mathfrak{A}$ (für große n) ist für viele Fragen der Analysis, z. B., wie Bohr gezeigt hat, für die Theorie der Dirichletschen Reihen (und erst recht für diejenige der Potenzreihen) von grundlegender Wichtigkeit. Wichtige dahin zielende Resultate hat Toeplitz angegeben. Etwas anders gerichtet ist eine Untersuchung von M. Riesz, der die Gesamtheit der zu beliebigen O. Hölderschen Exponenten gehörigen Bereiche betrachtet. Einen abstraktallgemeinen Ansatz über die hier sinngemäß überhaupt in Betracht kommenden Bildungen verdankt man E. Helly, der an die Minkowskische Geometrie und die *allgemeinen* O. Hölderschen Ungleichheiten über konvexe Funktionen anknüpft, zugrunde legt. Abschließende Ergebnisse über diese vielversprechenden Probleme, die eine Art Funktionalstatistik bezwecken, liegen freilich nicht vor.

§ 23. Treppenfunktionen.

Wir gehen nunmehr zu den Hermiteschen Matrizen im besonderen über und werden dabei die Hilbertschen analytischen Untersuchungen über unendliche Hermitesche Matrizen auf den hier vorliegenden alge-

braischen Fall übertragen, um den Überblick bei unendlichen Matrizen zu erleichtern und um für die Untersuchung der unendlichen Matrizen nur das übrigzulassen, was für die unendlichen Matrizen wirklich charakteristisch ist.

Es seien $f(\mu)$, $\sigma(\mu)$ zwei für $a \leqq \mu \leqq b$ erklärte, beschränkte, nicht notwendig reellwertige Funktionen und $\mu_0, \mu_1, \ldots, \mu_N, \mu_{N+1}$ etwa voneinander verschiedene Einteilungspunkte, so daß

$$a = \mu_0 < \mu_1 < \cdots < \mu_N < \mu_{N+1} = b$$

gilt. Es sei endlich ν_{q-1} ein beliebiger Punkt des q-ten abgeschlossenen Teilintervalles, also $\mu_{q-1} \leqq \nu_{q-1} \leqq \mu_q$. Man bilde die Summe

$$S = \sum_{q=1}^{N+1} f(\nu_{q-1})\,[\sigma(\mu_q) - \sigma(\mu_{q-1})]. \tag{61}$$

Setzen wir voraus, daß es eine Zahl A von der folgenden Beschaffenheit gibt: zu jedem $\varepsilon > 0$ kann ein $\eta > 0$ angetroffen werden derart, daß $|S - A| < \varepsilon$ wird, sobald nur die Länge $\mu_q - \mu_{q-1}$ eines jeden Teilintervalles $< \eta$ ist. Es soll also $|S - A| < \varepsilon$ für $\mu_q - \mu_{q-1} < \eta$ gelten, wie auch dabei die Wahl von N, der Einteilungspunkte μ_q und der Zwischenpunkte ν_q getroffen sein mag. Es kann offenbar höchstens eine von ε unabhängige Zahl A von dieser Beschaffenheit geben. Wenn es eine gibt, so sagen wir, der Grenzwert $\lim S = A$ sei vorhanden. Man bezeichnet diesen von Stieltjes eingeführten Grenzwert sinngemäß mit

$$\int_a^b f(\mu)\, d\sigma(\mu),$$

wobei $d\sigma(\mu)$ in einem anderen Sinne als bei der gewöhnlichen Integration als ein „Differential“ aufgefaßt werden soll (vgl. § 49). Die Funktion $\sigma(\mu)$ kann Sprünge haben und dann ist das Stieltjessche Integral eine unstetige Funktion von b. Ist $\lim S$ vorhanden, so sagen wir, die Funktion f sei auf $[a, b]$ in bezug auf die Dichte σ integrierbar. Im Spezialfall $\sigma(\mu) \equiv \mu$ handelt es sich um den Riemannschen Integralbegriff. Entsprechend erklärt man $\int_0^{+\infty}$ als $\lim\limits_{b \to +\infty} \int_0^b$ usf. Offenbar gilt (bei konstanten c)

$$\begin{aligned} &\int_a^b (c_1 f_1 + c_2 f_2)\, d\sigma = c_1 \int_a^b f_1\, d\sigma_1 + c_2 \int_a^b f_2\, d\sigma; \\ &\int_a^b f\, d(\sigma_1 + \sigma_2) = \int_a^b f\, d\sigma_1 + \int_a^b f\, d\sigma_2, \end{aligned} \tag{62}$$

sofern die Integrale rechterhand vorhanden sind, ferner

$$(63)\qquad \int_a^b f\,d\sigma = \int_a^\xi f\,d\sigma + \int_\xi^b f\,d\sigma,$$

wenn $a < \xi < b$ ist und das Integral linkerhand existiert. Da nämlich $\lim S$ unabhängig von der Art der bei den Näherungssummen S benutzten Einteilungen sein soll, so können wir $\lim S$ unter ausschließlicher Benutzung von Einteilungen berechnen, bei welchen ξ immer ein μ_q ist. — Addiert man zu $\sigma(\mu)$ eine Konstante, so ändert man dadurch den Wert des Integrals nicht ab.

Es seien $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(p)}$ endlich viele, der Größe nach geordnete verschiedene Zahlen auf der μ-Geraden und es mögen (nicht notwendig reelle) Zahlen $C_0, C_1, \ldots, C_p$ vorhandcn sein derart, daß die Funktion $\sigma(\mu)$ für

$$-\infty < \mu < \lambda^{(1)},\ \lambda^{(1)} \leqq \mu < \lambda^{(2)},\ \ldots,\ \lambda^{(p-1)} \leqq \mu < \lambda^{(p)},\ \lambda^{(p)} \leqq \mu < +\infty$$

bzw. identisch gleich

$$C_0,\qquad C_1,\qquad \ldots,\qquad C_{p-1},\qquad C_p$$

ist. Wir sagen dann, $\sigma(\mu)$ sei eine Treppenfunktion. Sie ist für alle reellen μ erklärt, kann höchstens $p+1$ Werte annehmen und erleidet im Punkte $\lambda^{(j)}$ einen Sprung, wenn nicht zufällig $C_{j-1} = C_j$ gilt. Man setze zur Abkürzung $\underset{\lambda}{\Delta}\sigma = \sigma(\lambda+0) - \sigma(\lambda-0)$. Wir werden in § 39 sehen, daß in bezug auf eine Treppenfunktion jede stetige Funktion f auf jedem endlichen oder unendlichen Intervall integrabel ist, und zwar hat man

$$(61\text{a})\qquad \int_{-\infty}^{+\infty} f(\mu)\,d\sigma(\mu) = \sum_{j=1}^{p} f(\lambda^{(j)})\,\underset{\lambda^{(j)}}{\Delta}\sigma$$

und allgemeiner

$$(61\text{b})\qquad \int_{-\infty}^{\mu} f(\mu)\,d\sigma(\mu) = \sum_{\lambda^{(j)} \leqq \mu} f(\lambda^{(j)})\,\underset{\lambda^{(j)}}{\Delta}\sigma,$$

wobei letztere Summation so zu verstehen ist, daß man über diejenigen j zu summieren hat, für welche $\lambda^{(j)} \leqq \mu$ gilt. Der Wert des Integrals hängt also bei festem σ nur von den Werten der Funktion ab, die zu $\mu = \lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(p)}$ gehören (für die übrigen μ braucht $f(\mu)$ eigentlich überhaupt nicht erklärt zu sein). Es ist also trivial, daß bei festem σ, also bei festem p

$$(61\text{c})\qquad \int_{-\infty}^{+\infty} \lim_{m=+\infty} f_m(\mu)\,d\sigma(\mu) = \lim_{m=+\infty} \int_{-\infty}^{+\infty} f_m(\mu)\,d\sigma(\mu)$$

gilt, wenn die Grenzfunktion vorhanden und stetig ist. Bei den endlichen Matrizen werden wir es nur mit Treppenbelegungen zu tun haben. Es handelt sich also um eine integrale, an spätere funktionale Zwecke (S. 167ff.) gleich hier angepaßte Schreibweise für Summen mit endlich vielen Summanden.

§ 24. Die Spektralmatrix.

Es sei $\mathfrak{A}$ eine Hermitesche Matrix mit den nicht notwendig verschiedenen Eigenwerten $\lambda_1, \lambda_2, \ldots, \lambda_n$, und $\mathfrak{U}$ eine dem $\mathfrak{A}$ derart angepaßte unitäre Matrix, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ diagonal ist, so daß $\mathfrak{A}$ durch $\mathfrak{U}$ in eine „Hauptachsenlage" „gedreht" wird. Man ordne der Matrix $\mathfrak{A}$ mittels $\mathfrak{U}$ und der Eigenwerte λ_j die n^2 Funktionen

$$\sigma_{ik}(\mu) = \sum_{\lambda_j \leqq \mu} \bar{u}_{ji} u_{jk} \tag{64}$$

zu. Die Summation ist in demselben Sinne wie bei (61b) zu verstehen. Nur hat man jetzt zu beachten, daß die λ_j nicht verschieden zu sein brauchen, so daß sowohl $\bar{u}_{j_1 i} u_{j_1 k}$ als auch $\bar{u}_{j_2 i} u_{j_2 k}$ in die Summation auch dann aufzunehmen ist, wenn $\lambda_{j_1} = \lambda_{j_2} \leqq \mu$ gilt (sonst sind die herzuleitenden Formeln bereits für $\mathfrak{A} = \mathfrak{E}$ falsch). Durch (64) wird eine Matrix $\| \sigma_{ik}(\mu) \|$ von Treppenfunktionen erklärt, die Spektralmatrix von $\mathfrak{A}$ heißt. Da die λ_j nicht der Größe nach numeriert zu sein brauchen[1]), so ist $\mathfrak{U}$ durch die Forderung, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ eine Diagonalmatrix sein soll, nicht einmal dann eindeutig bestimmt, wenn alle Eigenwerte verschieden sind[2]). Wir werden sehen, daß $\mathfrak{A}$ dennoch nur eine Spektralmatrix hat: die Unbestimmtheit für die unter (64) stehenden ***Funktionen*** der u hebt sich weg, wie auch dabei das Paar (i, k) fest gewählt sein mag. Bezeichnen, wie früher, $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(p)}$ die verschiedenen Zahlen unter den n Eigenwerten $(1 \leqq p \leqq n)$, so kann (64) in der Gestalt

$$\sigma_{ik}(\mu) = \sum_{\lambda^{(j)} \leqq \mu} \mathop{\Delta}_{\lambda^{(j)}} \sigma_{ik}; \quad \mathop{\Delta}_{\lambda^{(j)}} \sigma_{ik} = \sigma_{ik}(\lambda^{(j)} + 0) - \sigma_{ik}(\lambda^{(j)} - 0) \tag{65}$$

geschrieben werden und unsere Behauptung läuft darauf hinaus, daß durch $\mathfrak{A} = \| a_{ik} \|$ nicht nur die λ, sondern auch die Matrizen

$$\| \mathop{\Delta}_{\lambda^{(1)}} \sigma_{ik} \|, \ldots, \| \mathop{\Delta}_{\lambda^{(p)}} \sigma_{ik} \| \tag{66}$$

---

[1]) Man kann ein dreiachsiges Ellipsoid auf mehrere Weisen „in eine Hauptachsenlage drehen".

[2]) Ein Rotationsellipsoid kann man um die Rotationsache beliebig drehen, ohne es dabei in eine Lage zu bringen, die keine Hauptachsenlage ist.

vollkommen bestimmt sind (von den Matrizen

(67) $$\|\bar{u}_{1i} u_{1k}\|, \|\bar{u}_{2i} u_{2k}\|, \ldots, \|u_{ni} u_{nk}\|$$

wird dies nicht behauptet).

Es gilt nach Voraussetzung, mindestens bei einer bestimmten Numerierung der λ_i (und bei fest gegebener Numerierung der a_{ik}), die Beziehung $\|\lambda_i \delta_{ik}\| = \mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$, also $\mathfrak{A} = \mathfrak{U}^{-1}\|\lambda_i \delta_{ik}\|\mathfrak{U}$, mithin wegen $\mathfrak{U}^{-1}\mathfrak{E}\mathfrak{U} = \mathfrak{E}$ auch

$$(\lambda\mathfrak{E} - \mathfrak{A})^{-1} = (\mathfrak{U}^{-1}[\lambda\mathfrak{E} - \|\lambda_i \delta_{ik}\|]\mathfrak{U})^{-1} = \mathfrak{U}^{-1}\|(\lambda - \lambda_i)\delta_{ik}\|^{-1}\mathfrak{U},$$

also, da $\mathfrak{U}^{-1} = \mathfrak{U}^*$ und $\|(\lambda - \lambda_i)\delta_{ik}\|^{-1} = \left\|\frac{\delta_{ik}}{\lambda - \lambda_i}\right\|$ ist, offenbar

$$(\lambda\mathfrak{E} - \mathfrak{A})^{-1} = \left\|\sum_{j=1}^{n}\sum_{l=1}^{n} \bar{u}_{ji}\frac{\delta_{jl}}{\lambda - \lambda_j} u_{lk}\right\| = \left\|\sum_{j=1}^{n}\frac{\bar{u}_{ji} u_{jk}}{\lambda - \lambda_j}\right\|.$$

Hierfür kann man wegen (61a), (61b), (64), (65), wenn wieder $(\lambda\mathfrak{E} - \mathfrak{A})^{-1} = \|R_{ik}(\lambda)\|$ gesetzt wird,

(68) $$R_{ik}(\lambda) = \sum_{j=1}^{p}\frac{\underset{\lambda^{(j)}}{\Delta}\sigma_{ik}}{\lambda - \lambda^{(j)}} = \int_{-\infty}^{+\infty}\frac{d\sigma_{ik}(\mu)}{\lambda - \mu}$$

schreiben. Folglich ist $\underset{\lambda^{(j)}}{\Delta}\sigma_{ik}$ das Residuum von $R_{ik}(\lambda)$ für $\lambda = \lambda^{(j)}$. Da aber $\mathfrak{A}$ nur eine Resolvente hat, so ist damit gezeigt, daß es nur eine Spektralmatrix gibt.

Ist $|\lambda|$ größer als der Betrag des absolut größten Eigenwertes, so sind alle $R_{ik}(\lambda)$ regulär, und zwar auch für $\lambda = \infty$. Wir wissen dies übrigens auch von früher her (S. 17). Mit Rücksicht auf die Schlußbemerkung von § 23 können wir daher mit dem Integral (68) so wie mit endlichen Summen operieren. Wegen $\frac{1}{\lambda - \mu} = \sum_{n=0}^{\infty}\frac{\mu^n}{\lambda^{n+1}}$ und da $\int_{-\infty}^{+\infty} = \int_{-\varrho}^{\varrho}$ ist, wenn $\varrho > |\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|$ gewählt wird, gilt daher die für $|\lambda| > \varrho$ konvergente Entwicklung

$$(\lambda\mathfrak{E} - \mathfrak{A})^{-1} = \|R_{ik}(\lambda)\| = \left\|\sum_{n=0}^{\infty}\frac{\int_{-\infty}^{+\infty}\mu^n d\sigma_{ik}(\mu)}{\lambda^{n+1}}\right\| = \sum_{n=0}^{\infty}\frac{\left\|\int_{-\infty}^{+\infty}\mu^n d\sigma_{ik}(\mu)\right\|}{\lambda^{n+1}}.$$

Der Vergleich mit der für $|\lambda| > \varrho$ ebenfalls gültigen C. Neumannschen

Entwicklung (S. 19) ergibt, daß

$$(69) \qquad \mathfrak{A}^n = \left\| \int_{-\infty}^{+\infty} \mu^n \, d\sigma_{ik}(\mu) \right\|; \qquad n = 0, 1, 2, \ldots,$$

also insbesondere

$$(70) \qquad \delta_{ik} = \int_{-\infty}^{+\infty} d\sigma_{ik}(\mu), \qquad a_{ik} = \int_{-\infty}^{+\infty} \mu \, d\sigma_{ik}(\mu)$$

gilt. Mit Rücksicht auf den linearen Charakter der Integration folgt aus (69)

$$(71) \qquad \Phi(\mathfrak{A}) = \left\| \int_{-\infty}^{+\infty} \Phi(\mu) \, d\sigma_{ik}(\mu) \right\|,$$

wobei $\Phi(X)$ eine ganze rationale Funktion bezeichnet. Doch können wir diese Formel unter Beachtung des vorher Gesagten auch für alle Potenzreihen in Anspruch nehmen, deren Konvergenzgebiet die n Eigenwerte enthält. Mit $\Phi(X) = e^{tX}$ folgt so z. B. für die auf S. 19 eingeführte Matrizenexponentielle die Laplace-Fouriersche Darstellung

$$(72) \qquad \exp(\mathfrak{A} t) = \left\| \int_{-\infty}^{+\infty} e^{\mu t} \, d\sigma_{ik}(\mu) \right\|.$$

Aus (64) ist ersichtlich, daß $\sigma_{ik}(-\infty) = 0$, $\sigma_{ik}(+\infty) = \delta_{ik}$ ist. Für $\mu < -\varrho$ ist nämlich die Summation *leer*, für $\mu > \varrho$ gilt aber

$$\sigma_{ik}(\mu) = \sum_{j=1}^{n} \bar{u}_{ji} u_{jk}$$

und dies ist $= \delta_{ik}$, da $\mathfrak{U}$ unitär ist. Es gilt ferner aus demselben Grund

$$\sum_{\lambda_j \leqq \mu} |\bar{u}_{ji} u_{jk}| \leqq \sum_{j=1}^{n} |\bar{u}_{ji} u_{jk}| \leqq \sqrt{\sum_{j=1}^{n} |u_{ji}|^2} \sqrt{\sum_{k=1}^{n} |u_{jk}|^2} = \sqrt{1}\sqrt{1},$$

so daß die Summe der Beträge der Sprünge von $\sigma_{ik}(\mu)$, also wegen $\sigma_{ik}(-\infty) = 0$ auch $|\sigma_{ik}(\mu)|$ selbst, für alle μ gewiß $\leqq 1$ ist. Die Funktionen $\sigma_{ii}(\mu)$ sind, wegen (64) und da $\bar{u}_{ji} u_{ji} = |u_{ji}|^2$ ist, $\geqq 0$ und nicht abnehmend. Die übrigen $\sigma_{ik}(\mu)$ brauchen nicht reellwertig zu sein. Allerdings ist die Spektralmatrix bei jedem μ offenbar von Hermiteschem Typus, $\sigma_{ik}(\mu) = \bar{\sigma}_{ki}(\mu)$, wobei wir, wie auch künftig, $\bar{f}(\mu) = \overline{f(\mu)}$ setzen. Die Kopplungsform der Spektralmatrix von $\mathfrak{A}$ ist in jedem festen Punkte von E eine nicht abnehmende Funktion von μ, die für $\mu < -\varrho$ verschwindet und für $\mu > \varrho$ gleich 1 ist. Es ist

nämlich wegen (64)

$$
\begin{aligned}
(73)\quad \Phi(\|\sigma_{ik}(\mu)\|;\, x_1, x_2, \ldots, x_n) &= \sum_{i=1}^{n}\sum_{k=1}^{n}\sigma_{ik}(\mu)\, x_i \bar{x}_k \\
&= \sum_{\lambda_j \leqq \mu}\ \sum_{i=1}^{n}\bar{u}_{ji}x_i \sum_{k=1}^{n} u_{jk}\bar{x}_k = \sum_{\lambda_j\leqq\mu}\Big|\sum_{i=1}^{n}\bar{u}_{ji}x_i\Big|^2,
\end{aligned}
$$

was mit μ nicht abnimmt, für $\mu < -\varrho$, oder genauer für $\mu < \mathsf{n}(\mathfrak{A})$ die leere Summe ergibt, also $= 0$ ist, und für $\mu > \mathsf{m}(\mathfrak{A})$ in

$$
\sum_{j=1}^{n}\Big|\sum_{i=1}^{n}\bar{u}_{ji}x_i\Big|^2,
$$

also in die Kopplungsform der Norm einer unitären Matrix übergeht. Nun sind aber die Normen der unitären Matrizen $= \mathfrak{E}$, so daß die zugehörige Kopplungsform auf E identisch $= 1$ ist.

Die Spektralmatrix stellt in einem leicht verständlichen Sinne, der hier nicht näher erläutert werden soll, eine Verteilungsmatrix dar, welche die Statistik der Eigenwerte von $\mathfrak{A}$ beschreibt. Die erste Relation (70) sagt aus, daß die Gesamtwahrscheinlichkeit die Gewißheit ist. Wir werden jetzt eine Identität herleiten, welche die Unabhängigkeit der Wahrscheinlichkeiten der verschiedenen Spektralgebiete ausspricht und übrigens das elementare Analogon der Hellingerschen Spektraltheorie enthält (vgl. § 76).

Das Intervall $\alpha_1 \leqq \mu \leqq \alpha_2$ wollen wir mit a bezeichnen. Wir sprechen von einem Intervall auch dann, wenn es auf einen Punkt zusammenschrumpft und setzen, in Verallgemeinerung der auf S. 49 eingeführten Schreibweise, $\underset{\mathrm{a}}{\Delta}\sigma = \sigma(\alpha_2 + 0) - \sigma(\alpha_1 - 0)$, wobei $\alpha_2 \geqq \alpha_1$ ist. Unter dem Produkt a b zweier Intervalle verstehen wir den Durchschnitt von a und b, d. h. das Intervall, das aus allen gemeinsamen Punkten von a und b besteht. Z. B. ist a a = a, und a b ist ein Punkt, wenn a und b einen Endpunkt und sonst nichts gemein haben. Ist der Durchschnitt die leere Menge, so setzen wir a b = o und $\underset{\mathrm{o}}{\Delta}\sigma = 0$. Mit diesen Bezeichnungen kann der erwähnte, im nächsten Paragraph zu beweisende Unabhängigkeitssatz in der Gestalt

$$
(74)\qquad \|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\| \cdot \|\underset{\mathrm{b}}{\Delta}\sigma_{ik}\| = \|\underset{\mathrm{ab}}{\Delta}\sigma_{ik}\|
$$

geschrieben werden, wobei a, b ein beliebiges Paar von Intervallen bezeichnet. Nun ist die Reihe $\log(1-z) = \sum_{n=1}^{\infty}\frac{z^n}{n}$ für $|z| < 1$ konvergent (und übrigens auch für $|z| = 1$, wenn $z \neq 1$ ist, und stellt dabei die Funktion

dar). Folglich kann $\log \mathfrak{A}$ für jede Matrix erklärt werden, für die $[\mathsf{P}(\mathfrak{A})]^2 = \mathsf{M}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{M}(\mathfrak{A}^*\mathfrak{A}) < 1$ ausfällt (es ist dabei der Hauptzweig gemeint, für den $\log \mathfrak{E} = 2\pi\sqrt{-1}\, m\, \mathfrak{E} = \mathfrak{O}$, $m = 0$ ausfällt); vgl. übrigens (82) weiter unten. — Ist $\mathfrak{A}$ Hermitesch, so gilt $\mathfrak{A}\mathfrak{A}^* = \mathfrak{A}^2$. Andererseits ist $\mathsf{M}(\mathfrak{A}^2)$ der Betrag des absolut größten Eigenwertes von $\mathfrak{A}^2$, also $= [\mathsf{M}(\mathfrak{A})]^2$, da die Eigenwerte von $\mathfrak{A}^2$ die Quadrate der Eigenwerte von $\mathfrak{A}$ sind. Für Hermitesche $\mathfrak{A}$ geht also unsere Bedingung in $\mathsf{M}(\mathfrak{A}) < 1$ über. Nun ist aber $\|\sigma_{ik}(\mu)\|$, also auch $\|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\|$ immer Hermitesch. Aus $\alpha \leqq \beta$ folgt ferner, wie wir gesehen haben, $\mathsf{M}(\|\sigma_{ik}(\alpha)\|) \leqq \mathsf{M}(\|\sigma_{ik}(\beta)\|) \leqq 1$ und auch $\mathsf{M}(\|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\|) \leqq 1$, wobei das Gleichheitszeichen wegen (73) nur in einem uninteressanten Grenzfall gilt. So können wir von (74) zu

$$(75)\qquad \log\|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\| + \log\|\underset{\mathrm{b}}{\Delta}\sigma_{ik}\| = \log\|\underset{\mathrm{ab}}{\Delta}\sigma_{ik}\|$$

übergehen, da die Funktionalgleichung der Logarithmusfunktion aus der Maclaurinreihe gewonnen werden kann, also auch für *vertauschbare* Matrizenargumente gilt, und (75) ist nur eine andere Gestalt des Hellingerschen Unabhängigkeitssatzes.

§ 25. Fortsetzung. Die Spektralmatrix als Verteilungsmatrix.

Wir gehen nun zum Beweise der Unabhängigkeitsrelationen über. Es handelt sich dabei nur um eine Abzählung der Eigenwerte bzw. der „richtig normierten" Eigenlösungen, so daß dieser Paragraph überschlagen werden kann.

Wegen $\mathrm{ab} = \mathrm{ba}$ sind in (74) die beiden Faktoren linkerhand vertauschbar, und (74) ist symmetrisch. An Stelle von (74) genügt es, die Beziehung

$$(76)\qquad \|\sigma_{ik}(\nu)\|\,\|\sigma_{ik}(\mu)\| = \|\sigma_{ik}(\mu)\|\,\|\sigma_{ik}(\nu)\| = \|\sigma_{ik}(\mu)\|;\quad \mu \leqq \nu$$

zu beweisen. Wir zeigen vorerst, daß (74) in (76) enthalten ist. Sind α_1, α_2; β_1, β_2 die Endpunkte der Intervalle a; b, wobei $\alpha_1 \leqq \alpha_2$; $\beta_1 \leqq \beta_2$ vorausgesetzt ist, so gilt wegen $\underset{\mathrm{a}}{\Delta}\sigma_{ik} = \sigma_{ik}(\alpha_2 + 0) - \sigma_{ik}(\alpha_1 - 0)$:

$$(77)\qquad \begin{aligned} \|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\|\,\|\underset{\mathrm{b}}{\Delta}\sigma_{ik}\| &= \|\sigma_{ik}(\alpha_2+0)\|\,\|\sigma_{ik}(\beta_2+0)\| \\ &- \|\sigma_{ik}(\alpha_1-0)\|\,\|\sigma_{ik}(\beta_2+0)\| - \|\sigma_{ik}(\alpha_2+0)\|\,\|\sigma_{ik}(\beta_1-0)\| \\ &+ \|\sigma_{ik}(\alpha_1-0)\|\,\|\sigma_{ik}(\beta_1-0)\|, \end{aligned}$$

wobei die vier Produkte rechterhand mittels (76) ausgewertet werden können.

Wir betrachten zunächst den Fall, wo sowohl a als auch b ein Punkt, als $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$ ist. Um feste Vorstellungen zu haben, soll angenommen werden, daß a *vor* b liegt, so daß $\alpha_1 \leqq \beta_1$ gilt. Ist nun mindestens eine der beiden Punkte a, b kein Eigenwert, so ist mindestens einer der beiden Faktoren in (74) linkerhand gewiß die Nullmatrix, da $\underset{\mathrm{a}}{\Delta}\sigma_{ik} = \sigma_{ik}(\alpha_1 + 0) - \sigma_{ik}(\alpha_1 - 0)$ gilt und da $\sigma_{ik}(\mu)$ nur dann einen Sprung erleiden kann, wenn μ ein Eigenwert ist; vgl. (64). Aber auch die Matrix (74) rechterhand ist die Nullmatrix, da ab entweder die leere Menge oder aber ein Punkt ist, der kein Eigenwert ist (je nachdem $\mathrm{a} \neq \mathrm{b}$ oder $\mathrm{a} = \mathrm{b}$ gilt). Für diesen Fall ist also (74) trivial. — Es seien also a, b Punkte, die beide Eigenwerte sind. Wir betrachten dabei zunächst den Fall, wo a und b zusammenfallen, so daß $\lambda = \alpha_1 = \alpha_2 = \beta_1 = \beta_2$ gilt und daher (77) wegen (76) offenbar in

$$\|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\|\,\|\underset{\mathrm{b}}{\Delta}\sigma_{ik}\|$$
$$= \|\sigma_{ik}(\lambda + 0)\|^2 - 2\,\|\sigma_{ik}(\lambda + 0)\|\,\|\sigma_{ik}(\lambda - 0)\| + \|\sigma_{ik}(\lambda - 0)\|^2$$
$$= \|\sigma_{ik}(\lambda + 0)\| - 2\,\|\sigma_{ik}(\lambda - 0)\| + \|\sigma_{ik}(\lambda - 0)\|$$
$$= \|\sigma_{ik}(\lambda + 0)\| - \|\sigma_{ik}(\lambda - 0)\| = \|\underset{\lambda}{\Delta}\sigma_{ik}\|$$

übergeht. Nun ist aber $\lambda = \mathrm{a} = \mathrm{b}$, also $\lambda = \mathrm{ab}$, so daß (74) auch für diesen Fall verifiziert ist. Es seien endlich a, b Punkte, die beide Eigenwerte sind, aber nicht zusammenfallen: $\alpha_1 = \alpha_2 < \beta_1 = \beta_2$. Aus (77) bleibt wegen (76)

$$\|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\|\,\|\underset{\mathrm{b}}{\Delta}\sigma_{ik}\|$$
$$= \|\sigma_{ik}(\alpha_2 + 0)\| - \|\sigma_{ik}(\alpha_1 - 0)\| - \|\sigma_{ik}(\alpha_2 + 0)\| + \|\sigma_{ik}(\alpha_1 - 0)\|,$$

also die Nullmatrix. Aber auch $\|\underset{\mathrm{ab}}{\Delta}\sigma_{ik}\|$ ist die Nullmatrix, da die beiden Punkte a, b nicht zusammenfallen, also $\mathrm{ab} = \mathrm{o}$ ist. Damit ist gezeigt, daß (74) aus (76) gewiß folgt, wenn a und b Punkte und nicht eigentliche Intervalle sind. Wir haben dabei (76) auch für Grenzwertargumente $\mu \pm 0$ in Anspruch genommen. Dies war offenbar gestattet, da in Anbetracht dessen, daß die $\sigma_{ik}(\mu)$ Treppenfunktionen sind, $\sigma_{ik}(\mu \pm 0) = \sigma_{ik}(\mu \pm \varepsilon)$ gilt, wenn $\varepsilon\,(> 0)$ hinreichend klein ist.

Es seien nun a und b beliebig gelegene Intervalle (die auch Punkte sein können), $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(p)}$ die verschiedenen Eigenwerte von $\mathfrak{A}$, es möge ferner von ihnen a die Zahlen $\lambda_{\mathrm{a}}^{(1)}, \ldots, \lambda_{\mathrm{a}}^{(N)}$ und b die Zahlen

$\lambda_{\mathrm{b}}^{(1)}, \ldots, \lambda_{\mathrm{b}}^{(M)}$ enthalten, wobei sowohl N als auch $M \geqq 0$ und $\leqq p (\leqq n)$ ist. Es seien $\lambda_{\mathrm{ab}}^{(1)}, \ldots, \lambda_{\mathrm{ab}}^{(L)}$ die (voneinander verschiedenen) Zahlen, die in den beiden Folgen enthalten sind, d. h. die in ab liegenden Eigenwerte. Offenbar ist $\underset{\mathrm{a}}{\Delta} \sigma_{ik}$ die Summe der Sprünge, die zu den in a (evtl. am Rande von a) liegenden Eigenwerten gehören. Man hat daher

$$(78) \quad \|\underset{\mathrm{a}}{\Delta} \sigma_{ik}\| \, \|\underset{\mathrm{b}}{\Delta} \sigma_{ik}\| = \|\sum_{j=1}^{N} \underset{\lambda_{\mathrm{a}}^{(j)}}{\Delta} \sigma_{ik}\| \, \|\sum_{l=1}^{M} \underset{\lambda_{\mathrm{b}}^{(l)}}{\Delta} \sigma_{ik}\| = \sum_{j=1}^{N} \sum_{l=1}^{M} \|\underset{\lambda_{\mathrm{a}}^{(j)}}{\Delta} \sigma_{ik}\| \, \|\underset{\lambda_{\mathrm{b}}^{(l)}}{\Delta} \sigma_{ik}\| .$$

Nun ist das hinter dem Summenzeichen stehende Produkt die Nullmatrix oder $\|\underset{\lambda_{\mathrm{a}}^{(j)}}{\Delta} \sigma_{ik}\|$, je nachdem $\lambda_{\mathrm{a}}^{(j)} \neq$ oder $= \lambda_{\mathrm{b}}^{(l)}$ ist. Denn für Punkte ist (74) aus (76) bereits hergeleitet worden. Damit geht (78) über in

$$\|\underset{\mathrm{a}}{\Delta} \sigma_{ik}\| \, \|\underset{\mathrm{b}}{\Delta} \sigma_{ik}\| = \sum_{h=1}^{L} \|\underset{\lambda_{\mathrm{ab}}^{(h)}}{\Delta} \sigma_{ik}\| = \|\sum_{h=1}^{L} \underset{\lambda_{\mathrm{ab}}^{(h)}}{\Delta} \sigma_{ik}\| .$$

Dies ist aber wieder $= \|\underset{\mathrm{ab}}{\Delta} \sigma_{ik}\|$, da die $\lambda_{\mathrm{ab}}^{(h)}$ die in ab liegenden verschiedenen Eigenwerte bedeuten.

Damit ist gezeigt, daß (74) für alle Fälle gilt, wenn (76) richtig ist. Umgekehrt ist (76) in (74) enthalten. Denn die vorigen Verifikationen können, wie man leicht ersieht, auch in der umgekehrten Richtung gemacht werden. Es genügt dabei, wie wir gesehen haben, an Stelle von (74) den Spezialfall von (74) zu beweisen, wo beide Gebilde a, b Punkte sind. Ferner ist (74) für den Fall, wo mindestens einer der beiden Punkte a, b kein Eigenwert ist, eine Trivialität (siehe S. 54). Sind die beiden Punkte a, b Eigenwerte, so geht (74) in

$$(79) \qquad \|\underset{\lambda^{(j)}}{\Delta} \sigma_{ik}\| \, \|\underset{\lambda^{(l)}}{\Delta} \sigma_{ik}\| = \|\underset{\lambda^{(j)} \lambda^{(l)}}{\Delta} \sigma_{ik}\|$$

über. Für $j \neq l$ ist $\lambda^{(j)} \neq \lambda^{(l)}$, also $\lambda^{(j)} \lambda^{(l)} = 0$, mithin die Matrix rechterhand gleich $\|0\|$, während $\lambda^{(j)} \lambda^{(j)} = \lambda^{(j)}$ ist, so daß für (80) auch

$$(80) \qquad \|\underset{\lambda^{(j)}}{\Delta} \sigma_{ik}\| \, \|\underset{\lambda^{(l)}}{\Delta} \sigma_{ik}\| = \delta_{jl} \cdot \|\underset{\lambda^{(j)}}{\Delta} \sigma_{ik}\|$$

geschrieben werden kann. Wenn wir also (80) beweisen können, so wird zugleich (74) und damit auch (76) bewiesen sein. — Setzen wir nun die auf S. 50 gefundene Relation

$$(68)' \qquad R_{ik}(\lambda) = \sum_{j=1}^{p} \frac{\underset{j}{\Delta} \sigma_{ik}}{\lambda - \lambda^{(j)}} ,$$

wobei die Abkürzung $\underset{j}{\varDelta} = \underset{\lambda^{(j)}}{\varDelta}$ benutzt ist, in die Hilbertsche Funktionalgleichung (S. 16)

$$-(\lambda' - \lambda'')\,\|R_{ik}(\lambda')\|\,\|R_{ik}(\lambda'')\| = \|R_{ik}(\lambda')\| - \|R_{ik}(\lambda'')\|$$

oder
$$-(\lambda' - \lambda'')\sum_{r=1}^{n} R_{ir}(\lambda')\,R_{rk}(\lambda'') = R_{ik}(\lambda') - R_{ik}(\lambda'')$$

ein, so folgt

$$\sum_{r=1}^{n}\sum_{j=1}^{p}\sum_{l=1}^{p} \frac{\underset{j}{\varDelta}\sigma_{ir}}{\lambda' - \lambda^{(j)}}\,\frac{\underset{l}{\varDelta}\sigma_{rk}}{\lambda'' - \lambda^{(l)}}$$

$$= -\sum_{j=1}^{p} \underset{j}{\varDelta}\sigma_{ik}\left[\frac{1}{(\lambda'-\lambda'')(\lambda'-\lambda^{(j)})} - \frac{1}{(\lambda'-\lambda'')(\lambda''-\lambda^{(j)})}\right]$$

$$= \sum_{j=1}^{p} \frac{\underset{j}{\varDelta}\sigma_{ik}}{(\lambda' - \lambda^{(j)})(\lambda'' - \lambda^{(j)})},$$

also mit $\lambda = \lambda' = \lambda''$

$$(81)\qquad \sum_{j=1}^{p}\sum_{l=1}^{p} \frac{\sum\limits_{r=1}^{n} \underset{j}{\varDelta}\sigma_{ir}\,\underset{l}{\varDelta}\sigma_{rk}}{(\lambda - \lambda^{(j)})(\lambda - \lambda^{(l)})} = \sum_{j=1}^{p} \frac{\underset{j}{\varDelta}\sigma_{ik}}{(\lambda - \lambda^{(j)})(\lambda - \lambda^{(j)})},$$

und daher offenbar

$$(80)'\qquad \sum_{r=1}^{n} \underset{j}{\varDelta}\sigma_{ir}\,\underset{l}{\varDelta}\sigma_{rk} = \delta_{jl}\cdot\underset{j}{\varDelta}\sigma_{ik},$$

also auch (79). Damit ist alles bewiesen.

Übrigens kann (74) auch in der scheinbar mehr besagenden Gestalt

$$(82)\qquad \Big\|\int_{-\infty}^{+\infty} f(\mu)\,d\sigma_{ik}(\mu)\Big\|\,\Big\|\int_{-\infty}^{+\infty} g(\mu)\,d\sigma_{ik}(\mu)\Big\| = \Big\|\int_{-\infty}^{+\infty} f(\mu)\,g(\mu)\,d\sigma_{ik}(\mu)\Big\|$$

einer „Vollständigkeitsrelation“ geschrieben werden, wobei $f(\mu)$, $g(\mu)$ zwei willkürliche stetige Funktionen bezeichnen. Das Produkt linkerhand ist nämlich gleich

$$\Big\|\sum_{j=1}^{p} f(\lambda^{(j)})\underset{j}{\varDelta}\sigma_{ik}\Big\|\,\Big\|\sum_{j=1}^{p} g(\lambda^{(j)})\underset{j}{\varDelta}\sigma_{ik}\Big\| = \Big\|\sum_{r=1}^{n}\sum_{j=1}^{p} f(\lambda^{(j)})\underset{j}{\varDelta}\sigma_{ir}\sum_{l=1}^{p} g(\lambda^{(l)})\underset{l}{\varDelta}\sigma_{ik}\Big\|$$

$$= \sum_{j=1}^{p}\sum_{l=1}^{p} f(\lambda^{(j)})\,g(\lambda^{(l)})\,\Big\|\sum_{r=1}^{n}\underset{j}{\varDelta}\sigma_{ir}\underset{l}{\varDelta}\sigma_{rk}\Big\| = \sum_{j=1}^{p}\sum_{l=1}^{p} f(\lambda^{(j)})\,g(\lambda^{(l)})\,\|\delta_{jl}\underset{j}{\varDelta}\sigma_{ik}\|$$

$$= \Big\|\sum_{j=1}^{p}\sum_{l=1}^{p} f(\lambda^{(j)})\,g(\lambda^{(l)})\,\delta_{jl}\underset{j}{\varDelta}\sigma_{ik}\Big\| = \Big\|\sum_{j=1}^{p} f(\lambda^{(j)})\,g(\lambda^{(j)})\underset{j}{\varDelta}\sigma_{ik}\Big\|$$

$$= \Big\|\int_{-\infty}^{+\infty} f(\mu)\,g(\mu)\,d\sigma_{ik}(\mu)\Big\|.$$

§ 26. Einzelmatrizen und halbunitäre Matrizen.

Eine Hermitesche Matrix $\mathfrak{B}$, für welche $\mathfrak{B}^2 = \mathfrak{B}$ gilt, nennt man eine Einzelmatrix. Wegen (76) und (64) ist die Spektralmatrix einer jeden Hermiteschen Matrix $\mathfrak{A}$ eine (von μ abhängige) Einzelmatrix $\mathfrak{B}$. Die Transformationstheorie der Einzelmatrizen enthält darum die Transformationstheorie aller Hermiteschen Matrizen, indem sich die Matrix $\mathfrak{A}$ mit ihrer Spektralmatrix kogredient transformiert [und da $\|\sigma_{ik}(\mu)\|$ mit $\|\sigma_{ik}(\nu)\|$ wegen (76) vertauschbar ist[1])]. Wir wollen darauf jetzt nicht näher eingehen, da sich gegenüber den vorhergehenden nichts Neues ergeben würde. Hingegen ist bei unendlichen Matrizen, wie wir sehen werden, dieser Weg der allein gangbare.

Ist $\mathfrak{B}^2 = \mathfrak{B}$, so gilt auch $\mathfrak{B}^m = \mathfrak{B}$; $m > 0$, so daß die C. Neumannsche Reihe von $\mathfrak{B}$ in

$$\| R_{ik}(\lambda) \| = \sum_{m=0}^{\infty} \frac{\mathfrak{B}^m}{\lambda^{m+1}} = \frac{\mathfrak{B}^0}{\lambda} + \frac{\mathfrak{B}}{\lambda} \sum_{m=1}^{\infty} \frac{1}{\lambda^m} = \left\| \frac{\delta_{ik}}{\lambda} + \frac{b_{ik}}{\lambda(\lambda - 1)} \right\|$$

übergeht. Man überzeugt sich leicht, daß die letzte Darstellung von $R_{ik}(\lambda)$ auch dann gilt, wenn die Potenzreihe bereits divergiert. Die singulären Stellen, d. h. die Eigenwerte sind daher bei den Einzelmatrizen entweder $= 0$ oder $= 1$. Eine Zahl, die entweder $= 0$ oder $= 1$ ist, wollen wir mit γ_i bezeichnen. So können wir sagen, daß alle Einzelmatrizen unitär in eine Matrix $\|\gamma_i \delta_{ik}\|$ übergeführt werden können. Umgekehrt sind alle normalen Matrizen, deren Eigenwerte entweder $= 0$ oder $= 1$ sind, Einzelmatrizen. Aus $\mathfrak{B}^2 = \mathfrak{B}$ folgt nämlich zunächst, da $(\mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1})^2 = \mathfrak{U}\mathfrak{B}^2\mathfrak{U}^{-1} = \mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1}$ und $\mathfrak{B}$ Hermitesch ist, daß zugleich mit $\mathfrak{B}$ auch $\mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1}$ eine Einzelmatrix sein muß, so daß es genügt zu zeigen, daß $\|\gamma_i \delta_{ik}\|$ eine Einzelmatrix ist. Nun ist aber $\|\gamma_i \delta_{ik}\|^2 = \|\gamma_i^2 \delta_{ik}^2\| = \|\gamma_i \delta_{ik}\|$.

[Es kann also $\|\sigma_{ik}(\mu)\|$ mittels einer unitären Substitution $\mathfrak{U}$ bei jedem $\mathfrak{A}$ auf die Gestalt $\|\gamma_i \delta_{ik}\| = \mathfrak{U}\|\sigma_{ik}(\mu)\|\mathfrak{U}^{-1}$ gebracht werden, wobei $\mathfrak{U}$ in Anbetracht der Vertauschbarkeit von $\|\sigma_{ik}(\mu)\|$ mit $\|\sigma_{ik}(\nu)\|$ und mit Rücksicht auf die unter [1]), S. 31, gemachte Bemerkung als von μ unabhängig betrachtet werden kann. Es ist jedoch zu beachten, daß γ_i dennoch eine Funktion von μ ist.]

Da das Quadrat einer Hermiteschen Matrix gleich ihrer Norm ist und da die Normen bei jeder Matrix nichtnegativ definit sind, so ist jede Einzelmatrix, als Quadrat von sich selber (also von einer Hermiteschen Matrix), gewiß nichtnegativ definit.

---

[1]) Vgl. die Fußnote auf S. 31.

Eine (im allgemeinen nicht Hermitesche) Matrix $\mathfrak{G} = \| g_{ik} \|$ möge als eine von rechts halbunitäre Matrix genannt werden, wenn ihre hintere Norm $\mathfrak{G}\mathfrak{G}^*$ von der Gestalt $\| \gamma_i \delta_{ik} \|$ ist ($\gamma_i = 0$ oder $= 1$). Ist $\mathfrak{G}^*\mathfrak{G}$, die vordere Norm, von dieser Gestalt, so daß $\mathfrak{G}^*$ von rechts halbunitär ist, so nennen wir $\mathfrak{G}$ von links halbunitär. Es gibt Matrizen, die von rechts, aber nicht von links halbunitär sind. Die unitären Matrizen sind sowohl von rechts als auch von links halbunitär, und zwar mit $\gamma_1 = \gamma_2 = \cdots = \gamma_n = 1$. Die Matrizen von der Gestalt $\| \gamma_i \delta_{ik} \|$ sind ebenfalls sowohl von rechts als auch von links halbunitär. Die Kopplungsform der hinteren Norm einer von rechts halbunitären Matrix $\mathfrak{G} = \| g_{ik} \|$ ist wegen $\mathfrak{G}\mathfrak{G}^* = \| \gamma_i \delta_{ik} \|$ offenbar

$$= \sum_{i=1}^{n} \gamma_i | x_i |^2 \leqq \sum_{i=1}^{n} | x_i |^2 = 1.$$

Mit Rücksicht auf (53) ist daher auch $\mathsf{P}(\mathfrak{G}) \leqq 1$.

Die von rechts halbunitären Matrizen sind diejenigen, deren *hintere* Norm von der Gestalt $\| \gamma_i \delta_{ik} \|$, also gewiß eine Einzelmatrix ist. *Nicht jede* Einzelmatrix ist von der Gestalt $\| \gamma_i \delta_{ik} \|$. Wir zeigen nun, daß *jede* Einzelmatrix wohl als die *vordere* Norm einer passend zu wählenden, von rechts halbunitären Matrix dargestellt werden kann. Wenn man beachtet, daß jede Einzelmatrix, wenn auch nicht notwendig von der Gestalt $\| \gamma_i \delta_{ik} \|$ ist, aber gewiß in eine solche Matrix unitär übergeführt werden kann, so läßt sich der Satz in einigen Zeilen beweisen. Doch wäre diese Schlußweise, die den *fertigen* Satz über die unitäre Transformierbarkeit der Hermiteschen Matrizen auf die Diagonalgestalt heranzieht, auf unendliche Matrizen aus verschiedenen Gründen nicht übertragbar. Wir ziehen es daher vor, den Satz nach dem Vorgang von Hilbert mittels Jacobischer Transformation direkt zu beweisen, um dadurch Wiederholungen bei der Betrachtung unendlicher Matrizen zu vermeiden.

§ 27. Durchführung der Jacobischen Transformation.

Es sei $\mathfrak{B}$ eine Einzelmatrix. Wir haben zu beweisen, daß es eine von rechts halbunitäre Matrix $\mathfrak{G}$ gibt derart, daß $\mathfrak{B} = \mathfrak{G}^*\mathfrak{G}$ gilt. Die Matrix $\mathfrak{B}$ ist, wie erwähnt, nichtnegativ definit. Folglich sind darin nach der Formel (60) von § 21 alle Diagonalelemente $b_{ii} \geqq 0$. Sind alle $b_{ii} = 0$, so ist $\mathfrak{B}$ nach der erwähnten Formel die Nullmatrix, indem sodann auch die übrigen $n^2 - n$ Elemente verschwinden müssen. Diesen Fall können wir jedoch ausschließen, da für diesen Fall unsere Behauptung einfach mit $\mathfrak{G} = \| 0 \|$ bewiesen ist. Es möge also mindestens

ein b_{ii} geben, das > 0 ist. Um die Schreibweise zu vereinfachen, wollen wir annehmen, daß dies für $i = 1$ zutrifft, also $b_{11} > 0$ ist. Übrigens läuft dies nur auf die Vornahme einer (unitären) Permutationstransformation (§ 12) hinaus, womit gar nichts verändert wird. Die erste Zeile von $\|\mathfrak{G}\| = \|g_{ik}\|$ legen wir durch $g_{1i} = b_{11}^{-\frac{1}{2}} b_{1i}$ fest. Dann gilt, wegen $b_{11} > 0$ und da $\mathfrak{B}$ eine Einzelmatrix, also $\sum_{j=1}^{n} b_{1j} b_{j1} = b_{11}$ und $b_{1i} = \bar{b}_{i1}$ ist,

$$\sum_{j=1}^{n} g_{1j}\bar{g}_{1j} = \sum_{j=1}^{n} \frac{b_{1j}\bar{b}_{1j}}{b_{11}} = 1. \tag{83}$$

Wir setzen

$$^{(1)}b_{ik} = b_{ik} - \bar{g}_{1i} g_{1k} = b_{ik} - \frac{\bar{b}_{1i} b_{1k}}{b_{11}}. \tag{84}$$

Da $\mathfrak{B} = \|b_{ik}\|$ Hermitesch und b_{11} reell ist, so ist auch $^{(1)}\mathfrak{B} = \|^{(1)}b_{ik}\|$ Hermitesch. Unter Beachtung des Umstandes, daß $\mathfrak{B}$ eine Einzelmatrix ist, folgt, daß auch $^{(1)}\mathfrak{B}$ eine Einzelmatrix sein muß. Das (i, k)-te Element des Quadrates von $^{(1)}\mathfrak{B}$ ist nämlich

$$\sum_{j=1}^{n} {}^{(1)}b_{ij}\, {}^{(1)}b_{jk} = \sum_{j=1}^{n} \left(b_{ij} - \frac{\bar{b}_{1i} b_{1j}}{b_{11}}\right)\left(b_{jk} - \frac{\bar{b}_{1j} b_{1k}}{b_{11}}\right)$$

$$= \sum_{j=1}^{n} b_{ij} b_{jk} - \frac{b_{1k}}{b_{11}} \sum_{j=1}^{n} b_{ij} b_{j1} - \frac{b_{i1}}{b_{11}} \sum_{j=1}^{n} b_{1j} b_{jk} + \frac{b_{i1} b_{1k}}{b_{11}^2} \sum_{j=1}^{n} b_{1j} b_{j1}$$

$$= b_{ik} - \frac{b_{1k} b_{i1}}{b_{11}} - \frac{b_{i1} b_{1k}}{b_{11}} + \frac{b_{i1} b_{1k}}{b_{11}} = b_{ik} - \frac{\bar{b}_{1i} b_{1k}}{b_{11}} = {}^{(1)}b_{ik}.$$

Und zwar ist wegen (84) $^{(1)}b_{11} = 0$. Wenn $^{(1)}\mathfrak{B}$ nicht die Nullmatrix ist — auf die gegenteilige Annahme kommen wir später zurück —, so gibt es wieder ein i mit $^{(1)}b_{ii} > 0$. Und zwar muß $i > 1$ sein. Wir können wieder annehmen, daß bereits $^{(1)}b_{22} > 0$ ist. Wir legen dann die zweite Zeile von $\mathfrak{G}$ durch $g_{2i} = {}^{(1)}b_{22}^{-\frac{1}{2}}\, {}^{(1)}b_{2i}$ fest und setzen

$$^{(2)}b_{ik} = b_{ik} - \sum_{\nu=1}^{2} \bar{g}_{\nu i} g_{\nu k}. \tag{85}$$

Dann ist wieder $^{(2)}\mathfrak{B} = \|^{(2)}b_{ik}\|$ eine Einzelmatrix,

$$\sum_{j=1}^{n} g_{2j}\bar{g}_{2j} = 1 \quad \text{und dabei} \quad \sum_{j=1}^{n} g_{1i}\bar{g}_{2i} = 0. \tag{86}$$

Die erste dieser beiden Gleichungen beweist sich wie (83), die zweite gilt wegen

$$(87)\qquad \sum_{j=1}^{n} \bar{g}_{1j}\, g_{2j} = \sum_{j=1}^{n} \frac{\overline{b_{1j}}}{\sqrt{b_{11}}} \cdot \frac{b_{2j} - \dfrac{\bar{b}_{12}\, b_{1j}}{b_{11}}}{\sqrt{b_{22} - \dfrac{\bar{b}_{12}\, b_{12}}{b_{11}}}},$$

da diese Summe mit Rücksicht auf den Einzelcharakter von $\mathfrak{B}$ bis auf die beiden festen Nenner gleich

$$\sum_{j=1}^{n} \bar{b}_{1j}\left(b_{2j} - \frac{\bar{b}_{12}\, b_{1j}}{b_{11}}\right) = \sum_{j=1}^{n} b_{2j}\, b_{j1} - \frac{b_{21}}{b_{11}} \sum_{j=1}^{n} b_{1j}\, b_{j1} = b_{21} - \frac{b_{21}}{b_{11}}\, b_{11} = 0$$

ist. Es war vorher ${}^{(1)}b_{11} = 0$. Jetzt gilt, wie aus der Definition von ${}^{(2)}b_{ik}$ unmittelbar abgelesen werden kann, ${}^{(2)}b_{11} = {}^{(2)}b_{22} = 0$.

Setzen wir voraus, daß wir die q ersten Zeilen

$$(88)\qquad (g_{11}, \ldots, g_{1n}), \quad (g_{21}, \ldots, g_{2n}), \quad \ldots, \quad (g_{q1}, \ldots, g_{qn})$$

von $\mathfrak{G}$ derart erklärt haben, daß folgendes gilt: Es ist

$$(89)\qquad \sum_{j=1}^{n} g_{ij}\, \bar{g}_{kj} = \delta_{ik} \qquad \text{für } i = 1, 2, \ldots, q;\ k = 1, 2, \ldots, q,$$

und setzt man

$$(90)\qquad {}^{(q)}b_{ik} = b_{ik} - \sum_{\nu=1}^{q} \bar{g}_{\nu i}\, g_{\nu k},$$

so ist ${}^{(q)}\mathfrak{B} = \| {}^{(q)}b_{ik} \|$ eine Einzelmatrix; endlich gilt dabei ${}^{(q)}b_{\nu\nu} = 0$ für $\nu = 1, 2, \ldots, q$. — Wir können dann, wenn ${}^{(q)}\mathfrak{B}$ nicht die Nullmatrix ist, ebenso zeigen, daß man von q zu $q + 1$ übergehen kann, wie wir vorher von 1 zu 2 übergegangen sind. — Es sind nun zwei Fälle möglich.

Erstens ist es möglich, daß wir auf eine Zahl $q < n$ stoßen derart, daß ${}^{(q)}\mathfrak{B}$ die Nullmatrix ist. Wir besetzen sodann die $n - q$ unteren, bisher nicht erklärten Zeilen von $\mathfrak{G}$ mit lauter Nullen. Dann gilt wegen (89)

$$(91)\qquad \sum_{j=1}^{n} g_{ij}\, \bar{g}_{kj} = \gamma_i\, \delta_{ik} \qquad \text{für } i = 1, 2, \ldots, n;\ k = 1, 2, \ldots n,$$

wenn $\gamma_1 = \cdots = \gamma_q = 1,\ \gamma_{q+1} = \cdots = \gamma_n = 0$ gesetzt wird. Folglich ist $\mathfrak{G}$ von rechts halbunitär. Ferner gilt, wegen (90) und da ${}^{(q)}\mathfrak{B}$ die Nullmatrix ist,

$$(92)\qquad \| b_{ik} \| = \Big\| \sum_{\nu=1}^{q} \bar{g}_{\nu i}\, g_{\nu k} \Big\| = \Big\| \sum_{\nu=1}^{n} \bar{g}_{\nu i}\, g_{\nu k} \Big\|,$$

da ja $g_{\nu i} = 0$ ist für $\nu > q$. Es gibt daher eine Darstellung $\mathfrak{B} = \mathfrak{G}^* \mathfrak{G}$ von der gewünschten Art.

Zweitens ist es möglich, daß die Matrizen ${}^{(1)}\mathfrak{B}, {}^{(2)}\mathfrak{B}, \ldots, {}^{(n-1)}\mathfrak{B}$ erklärt werden können, indem dabei nicht einmal ${}^{(n-1)}\mathfrak{B}$ die Nullmatrix ist. Dann kann auch ${}^{(n)}\mathfrak{B}$ erklärt werden, und nach Voraussetzung ist ${}^{(n)}\mathfrak{B}$ eine Einzelmatrix mit ${}^{(n)}b_{\nu\nu} = 0$ für $\nu = 1, 2, \ldots, n$, also eine nichtnegativ definite Hermitesche Matrix mit durchweg verschwindenden diagonalen Elementen, d. h. nach der Formel (60) von § 21 die Nullmatrix. Für $q = n$ geht daher (90) unmittelbar in $\mathfrak{B} = \mathfrak{G}^* \mathfrak{G}$ über. Da nach Voraussetzung (90) gilt und jetzt $q = n$ ist, so ist $\mathfrak{G}$ von vornherein restlos erklärt und wegen (91) unitär, also erst recht von rechts halbunitär. — Damit ist gezeigt, daß jede Einzelmatrix als die vordere Norm einer von rechts halbunitären Matrix dargestellt werden kann.

Der Leser hat gewiß bemerkt, daß die vorhin benutzten, nach Jacobi benannten Transformationen, die zum Schluß zur „orthogonalen Kernzerspaltung", nämlich zu der Formel (92) führen, mit dem Orthogonalisierungsverfahren, auf das sich die Schursche Transformation gründet, im wesentlichen identisch sind.

§ 28. Die Jacobi-Toeplitzsche Parameterdarstellung der definiten Matrizen.

Ist $\mathfrak{T}$ eine kanonische Matrix, so ist ihre Determinante das Produkt der diagonalen Elemente, also dann und nur dann $\neq 0$, wenn keine von diesen verschwindet, d. h. wenn $\mathfrak{T}$ rekursiv ist (vgl. S. 11). In diesem und nur in diesem Falle sind wegen

$$\det(\mathfrak{T}\mathfrak{T}^*) = \det(\mathfrak{T}^*\mathfrak{T}) = (\det \mathfrak{T})(\det \mathfrak{T}^*) = |\det \mathfrak{T}|^2$$

die beiden Normen von $\mathfrak{T}$ nicht entartet, d. h. positiv definit und nicht nur nichtnegativ definit. Wir beweisen jetzt umgekehrt, daß jede positiv definite Hermitesche Matrix als Norm einer rekursiven Matrix dargestellt werden kann. Ob die hintere oder die vordere Norm gemeint ist, ist belanglos. Denn einerseits ist

$$(\mathfrak{T}\mathfrak{T}^*)^{-1} = \mathfrak{T}^{*-1}\mathfrak{T}^{-1} = \mathfrak{T}^{-1*}\mathfrak{T}^{-1};$$

andererseits sind $\mathfrak{T}$ und $\mathfrak{T}^{-1}$ zugleich rekursiv; endlich ist die Reziproke jeder positiv definiten Hermiteschen Matrix (mit Rücksicht auf das Verhalten der Eigenwerte beim Übergang zur reziproken Matrix) ebenfalls positiv definit.

Die Zerlegbarkeit der positiv definiten Hermiteschen Matrizen in das Produkt zweier Faktoren, die begleitende Matrizen voneinander sind und dabei einer der Faktoren kanonisch ist, ist seit Toeplitz für die Theorie der unendlichen Matrizen von grundlegender Wichtigkeit. — Der Beweis des Satzes kann wieder auf verschiedene, jedoch äquivalente Weisen geführt werden. Wir wählen den kürzesten Weg, der formale Wiederholungen des Vorangehenden scheinbar vermeidet.

Es sei $\mathfrak{A}$ eine positiv definite Hermitesche Matrix mit n^2 Elementen a_{ik}. Ist m irgendeine Zahl $< n$, und j eine der Zahlen $1, 2, \ldots, n$, so bezeichne $\mathfrak{A}_{[m;j]}$ die Matrix

$$(93)\qquad \mathfrak{A}_{[m;j]} = \left\| \begin{array}{cccc} a_{11} & a_{12} & a_{1\,m-1} & a_{1\,m} \\ a_{21} & a_{22} & a_{2\,m-1} & a_{2\,m} \\ & & & \\ & & & \\ a_{m-1\,1} & a_{m-1\,2} & a_{m-1\,m-1} & a_{m-1\,m} \\ a_{j\,1} & a_{j\,2} & a_{j\,m-1} & a_{j\,m} \end{array} \right\|,$$

die also nur m Zeilen hat. Die $m-1$ ersten Zeilen sind bez. aus den $m-1$ ersten Zeilen von $\mathfrak{A}$ unter Beibehaltung des Kolonnenzeigers entnommen, während die m-te Zeile auf dieselbe Weise aus der j-ten Zeile von $\mathfrak{A}$ entsteht. Es gilt, wegen $a_{ik} = \bar{a}_{ki}$ und da $\det \mathfrak{C} = \det \mathfrak{C}'$ und auch $= \overline{(\det \overline{\mathfrak{C}})}$ ist, offenbar

$$\overline{(\det \mathfrak{A}_{[m;j]})} = \det \mathfrak{A}'_{[m;j]}\,.$$

Setzt man in der Kopplungsform von $\mathfrak{A}$ die $n-m$ letzten Variablen gleich Null, so folgt, daß auf E auch die Kopplungsform von $\mathfrak{A}_{[m;m]}$ durchweg > 0 ist. Offenbar ist $\mathfrak{A}_{[m;m]}$ Hermitesch und positiv definit; denn wäre es nur nichtnegativ definit, so würde ein Eigenwert von $\mathfrak{A}_{[m;m]}$, also auch das Minimum von $\mathfrak{A}$ auf E und daher auch ein Eigenwert von $\mathfrak{A}$ verschwinden, während $\mathfrak{A}$ positiv definit ist. Es sind also alle Eigenwerte der Hermiteschen Matrix $\mathfrak{A}_{[m;m]}$ positiv, so daß

$$(94)\qquad \det \mathfrak{A}_{[m;m]} > 0$$

wird, was wegen $\mathfrak{A} = \mathfrak{A}_{[n;n]}$ auch für $m = n$ gilt; wir setzen $\det \mathfrak{A}_{[0;0]} = 1$. Dann gilt die im wesentlichen von Lagrange herrührende

Determinantenidentität

$$(95)\qquad a_{ik} = \sum_{j=1}^{n} \frac{\det \mathfrak{A}_{[j;i]}\,\overline{\det \mathfrak{A}_{[j;k]}}}{\det \mathfrak{A}_{[j-1;j-1]}\,\det \mathfrak{A}_{[j;j]}}; \qquad i, k = 1, 2, \ldots, n,$$

deren Verifizierung dem Leser überlassen bleiben möge. Setzt man nun

$$(96)\qquad \mathfrak{T} = \| t_{ik} \|; \qquad t_{ik} = \frac{\det \mathfrak{A}_{[k;i]}}{\sqrt{\det \mathfrak{A}_{[j-1;j-1]}\,\det \mathfrak{A}_{[j;j]}}},$$

so kann (95) unter Beachtung von (94), (96) in der Gestalt

$$(97)\qquad a_{ik} = \sum_{j=1}^{n} t_{ij}\,\bar{t}_{kj}$$

geschrieben werden, so daß $\mathfrak{A} = \mathfrak{T}\,\mathfrak{T}^*$ gilt. Für $m > j$ ist $\det \mathfrak{A}_{[m,j]} = 0$, da $\mathfrak{A}_{[m;j]}$ für $j < m$ nach Definition (93) zwei gleiche Zeilen hat. Wegen (96) gilt daher $t_{ik} = 0$ für $i < k$. Endlich ist t_{ii} wegen (94), (96) gewiß $\neq 0$. Mithin ist $\mathfrak{T}$ rekursiv und $\mathfrak{A} = \mathfrak{T}\,\mathfrak{T}^*$, w. z. b. w.

§ 29. Die Hermiteschen Matrizen von einfachem Spektrum und die Fourierschen Matrizen.

Die Matrix $\mathfrak{A}$ heißt von einfachem Spektrum, wenn alle ihre Eigenwerte verschieden sind, so daß mit unseren früheren Bezeichnungen $\lambda_j = \lambda^{(j)}$, $n = p$ gilt. Wir wollen in den folgenden Paragraphen die Hermiteschen Matrizen von einfachem Spektrum des näheren untersuchen. Die Formeln (64), (65), (70) gehen wegen $\lambda_j = \lambda^{(j)}$ in

$$(98)\qquad \begin{aligned} \sigma_{ik}(\mu) &= \sum_{\lambda_j \leqq \mu} \underset{j}{\Delta}\,\sigma_{ik}; & \underset{j}{\Delta}\,\sigma_{ik} &= \bar{u}_{ji}\,u_{jk}, \\ a_{ik} &= \sum_{j=1}^{n} \lambda_j \underset{j}{\Delta}\,\sigma_{ik}; & \delta_{ik} &= \sum_{j=1}^{n} \underset{j}{\Delta}\,\sigma_{ik} \end{aligned}$$

über, wobei zur Abkürzung $\underset{j}{\Delta} f = \underset{\lambda_j}{\Delta} f$ gesetzt ist.

Eine Treppenfunktion $\varrho(\mu)$, für welche

$$(99)\qquad \underset{j}{\Delta}\,\varrho = \underset{\lambda_j}{\Delta}\,\varrho > 0, \quad j = 1, 2, \ldots, n; \quad \underset{\mu}{\Delta}\,\varrho = 0, \ \mu \neq \lambda_j; \ \varrho(-\infty) = 0$$

gilt, nennen wir eine zu $\mathfrak{A}$ gehörige Dichtefunktion. Sie ist reellwertig, nicht abnehmend, erleidet an jedem Eigenwert einen Sprung und hat keine weiteren Sprungstellen, nimmt also genau $n+1$ Werte an. Zwei Dichtefunktionen, welche zu denselben λ_j, aber evtl. zu verschiedenen $\underset{j}{\Delta}\,\varrho$ gehören, also dieselben Sprungstellen, aber nicht notwendig dieselben Sprünge oder Wertevorräte haben, nennen wir äquivalent.

Wir sagen, die n Funktionen $\psi_1(\mu)$, $\psi_2(\mu)$, ..., $\psi_n(\mu)$ bilden ein zu $\varrho(\mu)$ gehöriges Orthogonalsystem, wenn

$$\int_{-\infty}^{+\infty} \psi_i(\mu)\,\overline{\psi}_k(\mu)\,d\varrho(\mu) = \delta_{ik}; \qquad i, k = 1, 2, \ldots, n \tag{100}$$

gilt. Wegen $\overline{\delta}_{ik} = \delta_{ik}$ bilden dann auch die konjugierten Funktionen $\overline{\psi}_i(\mu)$ ein zu $\varrho(\mu)$ gehöriges Orthogonalsystem, da ja das Integral gleich

$$\sum_{j=1}^{n} \psi_i(\lambda_j)\,\overline{\psi}_k(\lambda_j)\,\underset{j}{\Delta}\,\varrho = \delta_{ik} \tag{101}$$

und $\underset{j}{\Delta}\,\varrho$ reell ist. Zwei Funktionensysteme $\varphi_i(\mu)$, $\psi_i(\mu)$, für welche die n^2 Bedingungen $\psi_i(\lambda_j) = \varphi_i(\lambda_j)$ erfüllt sind, haben wir als nicht verschieden zu betrachten.

Wir wollen endlich die Matrix $\mathfrak{A}$ der n^2 Elemente a_{ik} als eine Fouriersche Matrix bezeichnen, wenn die Spektralmatrix eine Darstellung von der Form

$$\underset{\mu}{\Delta}\,\sigma_{ik} \equiv \psi_i(\mu)\,\overline{\psi}_k(\mu)\,\underset{\mu}{\Delta}\,\varrho; \qquad i, k = 1, 2, \ldots, n \tag{102}$$

gestattet, wobei $\varrho(\mu)$ eine passend wählbare Dichtefunktion mit genau n Sprungstellen und $\{\psi_i(\mu)\}$ irgendein zugehöriges Orthogonalsystem bedeutet. Aus (102) folgt mit Rücksicht auf die dritte Beziehung (98)

$$a_{ik} = \int_{-\infty}^{+\infty} \mu\,\psi_i(\mu)\,\overline{\psi}_k(\mu)\,d\varrho(\mu). \tag{102'}$$

Wir zeigen zunächst, daß jede Hermitesche Matrix $\mathfrak{A}$ von einfachem Spektrum gewiß eine Fouriersche Matrix ist. Es sei nämlich $\varrho(\mu)$ irgendeine der äquivalenten Belegungsfunktionen, deren Sprungstellen die Eigenwerte von $\mathfrak{A}$ sind, und man setze

$$\psi_i(\lambda_j) = \frac{\overline{u}_{ij}}{\sqrt{\underset{j}{\Delta}\,\varrho}}, \tag{103}$$

wobei die u dieselbe Bedeutung wie in (64) haben. Da für unsere Zwecke die $\psi_i(\mu)$ für $\mu \neq \lambda_j$ beliebige Werte annehmen können (vgl. S. 48), so können wir damit die $\psi_i(\mu)$ als erklärt betrachten; es sei $\psi_i(-\infty) = 0$. Wegen (98b), (103) gilt $\underset{j}{\Delta}\,\sigma_{ik} = \psi_i(\lambda_j)\,\overline{\psi}_k(\lambda_j)\,\underset{j}{\Delta}\,\varrho$. Doch ist (103) auch für $\mu \neq \lambda_j$ richtig, da für $\mu \neq \lambda_j$ beide Sprünge $\underset{\mu}{\Delta}\,\varrho$, $\underset{\mu}{\Delta}\,\sigma_{ik}$ verschwinden. Endlich bilden die $\psi_i(\mu)$ ein zu $\varrho(\mu)$ gehöriges Orthogonalsystem, da (101) oder (102) wegen (103), (98b) mit (98d) identisch sind. Mithin ist $\mathfrak{A}$, wie behauptet, gewiß eine Fouriersche Matrix.

Es wird sich herausstellen, daß umgekehrt jede Fouriersche Matrix, die offenbar von Hermiteschem Charakter ist [vgl. (102)'], von einfachem Spektrum sein muß.

Ist $\{\psi_i(\mu)\}$ ein zu $\varrho(\mu)$ gehöriges Orthogonalsystem, so ist die Matrix $\| \psi_i(\lambda_k) \sqrt{\underset{k}{\Delta} \varrho} \|$ wegen (101) unitär, also von nichtverschwindender Determinante. Wegen

$$\det \left\| \psi_i(\lambda_k) \sqrt{\underset{k}{\Delta} \varrho} \right\| = \sqrt{\underset{1}{\Delta} \varrho} \sqrt{\underset{2}{\Delta} \varrho} \cdots \sqrt{\underset{n}{\Delta} \varrho} \det \| \psi_i(\lambda_k) \|$$

und $\underset{j}{\Delta} \varrho > 0$ ist also auch $\det \| \psi_i(\lambda_k) \| \neq 0$. Mithin kann man zu jeder stetigen Funktion $f(\mu)$ auf genau eine Weise n Entwicklungskoeffizienten c_i derart bestimmen, daß

$$f(\mu) = \sum_{i=1}^{n} c_i \psi_i(\mu) \tag{104}$$

für $\mu = \lambda_1, \lambda_2, \ldots, \lambda_n$, also bei unseren Zwecken für alle μ gilt. Multipliziert man (104) mit $\overline{\psi}_k(\mu)$, so folgt unter Beachtung von (100), daß die Entwicklungskoeffizienten die Werte

$$c_i = \int_{-\infty}^{+\infty} f(\mu) \overline{\psi}_i(\mu) \, d\varrho(\mu)$$

haben, also sich auf die Fouriersche Weise bestimmen. Ist

$$g(\mu) = \sum_{i=1}^{n} d_i \overline{\psi}_i(\mu); \qquad d_i = \int_{-\infty}^{+\infty} g(\mu) \psi_i(\mu) \, d\varrho(\mu)$$

eine andere Funktion, so gilt wegen

$$\sum_{j=1}^{n} f(\lambda_j) g(\lambda_j) \underset{j}{\Delta} \varrho = \sum_{j=1}^{n} \sum_{i=1}^{n} c_i \psi_i(\lambda_j) \sum_{k=1}^{n} d_k \overline{\psi}_k(\lambda_j) \underset{j}{\Delta} \varrho$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{n} c_i d_k \sum_{j=1}^{n} \psi_i(\lambda_j) \overline{\psi}_k(\lambda_j) \underset{j}{\Delta} \varrho = \sum_{i=1}^{n} \sum_{k=1}^{n} c_i d_k \delta_{ik} = \sum_{j=1}^{n} c_i d_i$$

die Vollständigkeitsrelation

$$\int_{-\infty}^{+\infty} f(\mu) g(\mu) \, d\varrho(\mu) = \sum_{i=1}^{n} \int_{-\infty}^{+\infty} f(\mu) \overline{\psi}_i(\mu) \, d\varrho(\mu) \cdot \int_{-\infty}^{+\infty} g(\mu) \psi_i(\mu) \, d\varrho(\mu). \tag{105}$$

Es sei $F(\mu)$ eine beliebige stetige Funktion. Man setze

$$\mathfrak{F} = \left\| \int_{-\infty}^{+\infty} F(\mu) \varphi_i(\mu) \overline{\varphi}_k(\mu) \, d\varrho(\mu) \right\|, \quad \mathfrak{U} = \left\| \int_{-\infty}^{+\infty} \psi_i(\mu) \overline{\varphi}_k(\mu) \, d\varrho(\mu) \right\|, \tag{106}$$

wobei die $\varphi_i(\mu)$ ein von den $\psi_i(\mu)$ eventuell verschiedenes Orthogonalsystem bilden, aber zur selben Dichtefunktion $\varrho(\mu)$ wie die $\psi_i(\mu)$ gehören sollen. Mit wiederholter Benutzung der Vollständigkeitsrelation (105) folgt

$$\mathfrak{U}\mathfrak{F} = \left\| \sum_{j=1}^{n} \int_{-\infty}^{+\infty} \psi_i(\mu)\,\overline{\varphi}_j(\mu)\, d\varrho(\mu) \cdot \int_{-\infty}^{+\infty} F(\mu)\,\varphi_j(\mu)\,\overline{\varphi}_k(\mu)\, d\varrho(\mu) \right\|$$

$$= \left\| \int_{-\infty}^{+\infty} F(\mu)\,\psi_i(\mu)\,\overline{\varphi}_k(\mu)\, d\varrho(\mu) \right\|,$$

$$(\mathfrak{U}\mathfrak{F})\,\mathfrak{U}^* = \left\| \sum_{j=1}^{n} \int_{-\infty}^{+\infty} F(\mu)\,\psi_i(\mu)\,\overline{\varphi}_j(\mu)\, d\varrho(\mu) \cdot \int_{-\infty}^{+\infty} \overline{\psi}_k(\mu)\,\varphi_j(\mu)\, d\varrho(\mu) \right\|$$

$$= \left\| \int_{-\infty}^{+\infty} F(\mu)\,\psi_i(\mu)\,\overline{\psi}_k(\mu)\, d\varrho(\mu) \right\|,$$

oder ausführlicher geschrieben

$$(107)\quad \mathfrak{U} \left\| \int_{-\infty}^{+\infty} F(\mu)\,\varphi_i(\mu)\,\overline{\varphi}_k(\mu)\, d\varrho(\mu) \right\| \mathfrak{U}^* = \left\| \int_{-\infty}^{+\infty} F(\mu)\,\psi_i(\mu)\,\overline{\psi}_k(\mu)\, d\varrho(\mu) \right\|.$$

Setzt man hierin $F(\mu) \equiv 1$, so bleibt $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$, d. h. $\mathfrak{U}$ ist unitär[1]). Setzt man daher $F(\mu) \equiv \mu$, so folgt in Anbetracht der Relationen (102)', (106), (107) unmittelbar, daß zwei Fouriersche Matrizen gewiß unitär ineinander übergeführt werden können, wenn ihre Spektralmatrizen zur selben Dichtefunktion gehören [vgl. (102)]. Und zwar werden sie ineinander durch eine Matrix $\mathfrak{U}$ transformiert, die wegen (106) aus den Fourierschen Entwicklungskoeffizienten des einen Orthogonalsystems in bezug auf das andere gebildet wird.

§ 30. Fortsetzung. Orthogonale Polynome.

Das zu $\varrho(\mu)$ gehörige Orthogonalsystem $\{\varphi_i(\mu)\}$ möge ein Polynomsystem genannt werden, wenn darin $\varphi_i(\mu)$ von der Gestalt $\sum_{\nu=0}^{i-1} C_\nu^{(i)} \mu^\nu$, also eine ganze rationale Funktion *höchstens* vom Grade $i-1$ ist;

---

[1]) Der kundige Leser möge hierzu die Euler-Hurwitzsche Produktzerlegung der unitären Matrizen vergleichen (siehe S. 45). Sie versagt bei beliebigen unendlichen Matrizen, da es dann bei der Reduktion keinen *letzten* Schritt gibt oder eigentlich bereits darum, da eine unendliche normale Matrix u. U. nicht in eine Diagonalmatrix übergeführt werden kann. Hingegen kann die Hellingersche Formel (107), welche die unitäre Matrix aus infinitesimalen spektralen Drehungen (vgl. S. 160 ff.) zusammensetzt, auf unendliche Matrizen wohl übertragen werden. Vgl. insb. S. 187. Die sinngemäße Parameterdarstellung *aller* unendlichen unitären Matrizen findet man in § 101.

$i = 1, 2, \ldots, n$. Insbesondere ist dann $\varphi_1(\mu)$ eine Konstante, die wegen $\int\limits_{-\infty}^{+\infty} |\psi_1|^2 \, d\varrho = \int\limits_{-\infty}^{+\infty} \psi_1 \overline{\psi}_1 \, d\varrho = \delta_{11} = 1$ und $\underset{\mu}{\Delta}\, \varrho \geqq 0$ nicht verschwinden kann. Allgemeiner ist $C_{i-1}^{(i)} \neq 0$, also $\varphi_i(\mu)$ ein Polynom *genau* vom Grade $i - 1$. Denn nehmen wir an, daß es ein i mit $C_{i-1}^{(i)} = 0$ gibt, und es sei i_0 der kleinste derartige Zeiger. Es ist also $\varphi_j(\mu)$ ein Polynom genau vom Grade $j - 1$, wenn $j < i_0$ ist. Offenbar kann man daher das Monom μ^{j-1} aus den j ersten $\varphi(\mu)$ homogen-linear kombinieren. Man erkennt dies auf eine rekursive Weise, durch vollständige Induktion. — Da sich jedes Polynom vom Grade $\leqq j - 1$ aus den j ersten Potenzen, nämlich aus $\mu^0, \mu^1, \ldots, \mu^{j-1}$ aufbaut, so kann man allgemeiner jedes Polynom $\underset{j-1}{\Phi}(\mu)$ vom Grade $\leqq j - 1$ in der Gestalt

$$\underset{j-1}{\Phi}(\mu) = \sum_{\nu=1}^{j} A_\nu \varphi_\nu(\mu) \tag{108}$$

darstellen. Dies gilt nach Voraussetzung für $j < i_0$, also auch für $j = i_0 - 1$. Da φ_{i_0} nach unserer (zu widerlegenden) Annahme nicht vom genauen Grade $i_0 - 1$, sondern höchstens vom Grade $i_0 - 2 = j - 1$ ist, so müßte (108) für $\underset{j-1}{\Phi}(\mu) = \varphi_{i_0}(\mu)$ mit $j = i_0 - 1$ gelten, also

$$\int\limits_{-\infty}^{+\infty} \varphi_{i_0}(\mu)\, \overline{\varphi}_{i_0}(\mu)\, d\varrho(\mu) = \sum_{\nu=1}^{i_0-1} A_\nu \int\limits_{-\infty}^{+\infty} \varphi_\nu(\mu)\, \overline{\varphi}_{i_0}(\mu)\, d\varrho(\mu) \tag{109}$$

sein. Hierbei ist jedes $\nu < i_0$, also $\neq i_0$, so daß die rechte Seite wegen

$$\int\limits_{-\infty}^{+\infty} \varphi_i(\mu)\, \overline{\varphi}_k(\mu)\, d\varrho(\mu) = \delta_{ik} \tag{110}$$

verschwindet, während die linke Seite wegen (110) gleich 1 ist. — Mithin ist jedes $\varphi_i(\mu)$ genau vom Grade $i - 1$.

Folglich gilt (108) für $j = 1, 2, \ldots, n$. Ist $j < i \leqq n$, so folgt daher aus (108) mit Rücksicht auf (110)

$$\int\limits_{-\infty}^{+\infty} \underset{j-1}{\Phi}(\mu)\, \overline{\varphi}_i(\mu)\, d\varrho(\mu) = 0. \tag{111}$$

Hingegen kann (111) für *kein* Polynom $\underset{i-1}{\Phi}$ vom genauen Grade $i - 1$ richtig sein. Da dann nämlich

$$\underset{i-1}{\Phi} = C\mu^{i-1} + \underset{i-2}{\Phi} \quad \text{und} \quad C \neq 0$$

ist, so müßte (111) auch für $\underset{j-1}{\Phi} = \mu^{i-1}$ richtig sein, da (111) für

$\underset{j}{\Phi}_1 = \underset{i-2}{\Phi}$ gewiß richtig ist. Also müßte dann (111) für *jedes* Polynom Φ vom Grade $i-1$, also auch für $\Phi = \varphi_i$ richtig sein, während sich dabei $\delta_{ii} = 1 \neq 0$ ergeben muß. Folglich gilt (111) gewiß für $j \leqq i-1$, und es gilt gewiß nicht, wenn $\underset{j-1}{\Phi} = \underset{i-1}{\Phi}$ ist, ohne dabei ein Polynom vom Grade $< i-1$ zu sein. Dies alles gilt freilich auch dann, wenn $\bar{\varphi}_i$ durch φ_i ersetzt wird. Setzt man $\Phi = \mu\,\varphi_k(\mu)$, so folgt daher, daß das Integral

$$a_{ik} = \int_{-\infty}^{+\infty} \mu\,\varphi_i(\mu)\,\bar{\varphi}_k(\mu)\,d\varrho(\mu) \tag{112}$$

gewiß verschwindet, wenn die Differenz $i-k>1$ oder wenn $k-i>1$ ist, und daß es für $i-k=\pm 1$ gewiß nicht verschwindet. Nun ist aber (112) wegen (102), (110) das (i,k)-te Element einer Fourierschen Matrix, deren Spektralmatrix zu einem Polynomsystem gehört. Bei diesen Fourierschen Matrizen ist also

$$a_{ik} = 0 \quad \text{für } |i-k|>1; \qquad a_{i\,i+1} = \bar{a}_{i+1\,i} \neq 0; \qquad a_{ii} \gtreqless 0. \tag{113}$$

Es sei $\varrho(\mu)$ eine Dichtefunktion mit genau n voneinander verschiedenen Sprungstellen $\lambda_1, \lambda_2, \ldots, \lambda_n$. Da $\det \| \lambda_k^{i-1} \|$, als Vandermondesche Determinante voneinander verschiedener Zahlen, $\neq 0$, und auch $\underset{\lambda_k}{\Delta}\varrho = \underset{k}{\Delta}\varrho \neq 0$ ist (da ϱ für $\mu = \lambda_k$ einen Sprung erleidet), so ist auch

$$\det \left\| \sqrt{\underset{k}{\Delta}\varrho}\,\lambda_k^{i-1} \right\| = \sqrt{\underset{1}{\Delta}\varrho}\,\sqrt{\underset{2}{\Delta}\varrho}\ldots\sqrt{\underset{n}{\Delta}\varrho}\,\det \left\| \lambda_k^{i-1} \right\| \neq 0,$$

so daß die n Vektoren

$$\mathfrak{v}_\nu = \left(\sqrt{\underset{1}{\Delta}\varrho}\,\lambda_1^{\nu-1}, \sqrt{\underset{2}{\Delta}\varrho}\,\lambda_2^{\nu-1}, \ldots, \sqrt{\underset{n}{\Delta}\varrho}\,\lambda_n^{\nu-1}\right); \qquad \nu = 0, 1, \ldots, n-1$$

linear unabhängig sind. Nach dem Schmidtschen Verfahren folgt daraus die Existenz eines rekursiven Koeffizientensystems $C_{\nu-1}^{(i)}$ derart, daß, wenn $\mathfrak{u}_i = \sum_{\nu=1}^{i} C_{\nu-1}^{(i)}\,\mathfrak{v}_\nu$ gesetzt wird, $\mathfrak{u}_i\bar{\mathfrak{u}}_k = \delta_{ik}$; $i, k = 1, 2, \ldots, n$ gilt; siehe § 6. Setzt man daher $\varphi_i(\mu) = \sum_{\nu=1}^{i} C_{\nu-1}^{(i)}\,\mu^{\nu-1}$, so gilt

$$\int_{-\infty}^{+\infty} \varphi_i(\mu)\,\bar{\varphi}_k(\mu)\,d\varrho(\mu) = \sum_{j=1}^{n} \varphi_i(\lambda_j)\,\bar{\varphi}_k(\lambda_j)\,\underset{j}{\Delta}\varrho$$

$$= \sum_{j=1}^{n}\sum_{\nu=1}^{i} C_{\nu-1}^{(i)}\,\lambda_j^{\nu-1} \sum_{\tau=1}^{k} \bar{C}_{\tau-1}^{(k)}\,\lambda_j^{\tau-1}\,\underset{j}{\Delta}\varrho = \sum_{\nu=1}^{i}\sum_{\tau=1}^{k} C_{\nu-1}^{(i)}\,\bar{C}_{\tau-1}^{(k)}\,\mathfrak{v}_\nu\bar{\mathfrak{v}}_\tau = \mathfrak{u}_i\bar{\mathfrak{u}}_k = \delta_{ik}.$$

Es kann also zu jeder Dichtefunktion $\varrho(\mu)$ ein Polynomsystem $\varphi_1, \varphi_2, \ldots, \varphi_n$ angegeben werden, das zu $\varrho(\mu)$ als orthogonales Polynomsystem gehört.

§ 31. Fortsetzung. Jacobische Matrizen.

Eine Matrix $\mathfrak{A}$, welche den Bedingungen (113) genügt, nennt man eine Jacobische Matrix. Wir haben gesehen, daß jede Fouriersche Matrix, die zu einem orthogonalen Polynomsystem gehört, gewiß eine Jacobische Matrix ist. Wir werden jetzt nach E. Heine beweisen, daß umgekehrt jede Jacobische Matrix eine Fouriersche Matrix ist, deren Spektralmatrix sich aus einem orthogonalen Polynomsystem ableitet, und daß jede Jacobische Matrix von einfachem Spektrum (und offenbar von Hermiteschem Typus) ist.

Bevor wir zum Beweis übergehen, wollen wir einige Folgerungen erwähnen. Es sei vorgelegt irgendeine Fouriersche Matrix $\mathfrak{A}$. Es sei $\varrho(\mu)$ irgendeine der äquivalenten Dichtefunktionen, die zu $\mathfrak{A}$ gehören. Wir können sodann, wie wir vorhin bewiesen haben, ein orthogonales Polynomsystem bestimmen, das zu diesem $\varrho(\mu)$ gehört. Es bezeichne $\mathfrak{B}$ die Fouriersche Matrix, deren Spektralmatrix durch diese Polynome und durch $\varrho(\mu)$ bestimmt wird. $\mathfrak{A}$ und $\mathfrak{B}$ sind Fouriersche Matrizen, die dieselbe Dichtefunktion zulassen; sie sind daher, mit Rücksicht auf das bei (107) Gesagte, unitär äquivalent. Nun ist aber $\mathfrak{B}$ eine Jacobische Matrix, also von einfachem Spektrum. Mithin ist auch $\mathfrak{A}$ von einfachem Spektrum, so daß der Satz, wonach die Hermiteschen Matrizen von einfachem Spektrum Fouriersche Matrizen sind, umgekehrt werden kann, wie bereits auf S. 65 angegeben worden ist.

Zwei Fouriersche Matrizen sind dann und nur dann unitär äquivalent, wenn sie Dichtefunktionen zulassen, die in dem auf S. 63 erklärten Sinne äquivalent sind. Letzteres ist nämlich die notwendige und hinreichende Bedingung für die Übereinstimmung der Spektra. Für die unitäre Äquivalenz ist also nicht notwendig, daß beide Dichtefunktionen identisch sind. Dies folgt bereits daraus, daß die „Hellingersche Zerlegung“ (vgl. weiter unten) von $\|\sigma_{ik}(\mu)\|$ in einen von $\varrho(\mu)$ und in einen von dem Orthogonalsystem abhängigen Bestandteil vermöge (102) nicht eindeutig bestimmt ist, indem wir $\varrho(\mu)$ auf S. 64 innerhalb der äquivalenten Klasse willkürlich wählen konnten.

Die Kopplungsform der Sprungmatrix der Spektralmatrix einer Fourierschen Matrix läßt wegen (102) eine Darstellung von der Gestalt

$$\sum_{i=1}^{n}\sum_{k=1}^{n}\underset{\mu}{\Delta}\,\sigma_{ik}(\mu)\,x_i\bar{x}_k = \Big|\sum_{i=1}^{n}x_i\psi_i(\mu)\Big|^2\,\underset{\mu}{\Delta}\,\varrho$$

zu, während die Kopplungsform der Fourierschen Matrix wegen (98), (101), (102) gleich

$$\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}x_i\bar{x}_k = \int_{-\infty}^{+\infty}\mu\,\Big|\sum_{i=1}^{n}\psi_i(\mu)x_i\Big|^2\,d\varrho(\mu)$$

ist (und sie ist dann und nur dann negativer Werte fähig ist, wenn die nicht abnehmende Funktion ϱ mindestens eine negative Sprungstelle hat). Wir werden später sehen, wie diese Formeln von Hilbert und Hellinger auf Streckenspektra übertragen worden sind.

§ 32. Fortsetzung. Die Heineschen Formeln.

Wir gehen nun zu dem Beweis für die Behauptungen von § 31 über. Wegen (113) ist eine Diagonalmatrix gewiß keine Jacobische Matrix. Im binären Falle $n = 2$ ist die Matrix dann und nur dann eine Jacobische, wenn sie die Gestalt

$$\begin{Vmatrix} a_{11} & a_{12} \\ \bar{a}_{12} & a_{22} \end{Vmatrix}; \quad a_{12} \neq 0, \quad a_{11} \gtrless 0, \quad a_{22} \gtrless 0$$

hat. Im Falle $n = 1$ können wir der Einheitlichkeit halber etwa *jede* reelle „Matrix" als eine Jacobische gelten lassen. Doch wollen wir von jetzt an $n > 1$ voraussetzen. Setzen wir der Bequemlichkeit halber $a_{ii} = a_i$, $a_{i\,i+1} = -b_i$, so sind die Jacobischen Matrizen nach der Definition (113) von der Struktur

$$(114)\qquad \mathfrak{A} = \begin{Vmatrix} a_1 & -b_1 & & & & \\ -\bar{b}_1 & a_2 & -b_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\bar{b}_{n-2} & a_{n-1} & -b_{n-1} \\ & & & & -\bar{b}_{n-1} & a_n \end{Vmatrix}; \quad a_{ii} \gtrless 0, \quad b_i \neq 0,$$

wobei außerhalb der Diagonale und den beiden sie flankierenden Strahlen alle Elemente verschwinden. Die zu $\lambda \mathfrak{E} - \mathfrak{A}$ gehörigen homogenen Gleichungen sind daher

$$(115)\qquad \begin{aligned} &(\lambda - a_1) x_1 + b_1 x_2 = 0, \\ &\bar{b}_1 x_1 + (\lambda - a_2) x_2 + b_2 x_3 = 0, \\ &\bar{b}_2 x_2 + (\lambda - a_3) x_3 + b_3 x_4 = 0, \\ &\quad \cdots \\ &\bar{b}_{n-1} x_{n-1} + (\lambda - a_n) x_n = 0. \end{aligned}$$

Diese Gleichungen können bei jedem λ aufgelöst werden, nötigenfalls — wenn λ kein Eigenwert ist — durch $x_1 = \ldots = x_n = 0$. Es sei $x_i = x_i^0$ eine Lösung. Nach der ersten Gleichung und wegen $b_1 \neq 0$

ist $x_2^0 = x_1^0 \chi_2(\lambda)$, wobei $\chi_2(\lambda)$ ein Polynom vom genauen Grade 1 ist. Setzt man dies in die zweite Gleichung ein, so folgt wegen $b_2 \neq 0$, daß $x_3^0 = x_1^0 \chi_3(\lambda)$ gilt, wobei $\chi_3(\lambda)$ ein Polynom vom genauen Grade 2 bezeichnet. So kann man weitergehen. Aus der $(n-1)$-ten (nicht angeschriebenen) Gleichung folgt $x_n^0 = x_1^0 \chi_n(\lambda)$. Die n-te Gleichung nimmt daher eine Sonderstellung ein, sie wird durch das Wertepaar x_{n-1}^0, x_n^0 entweder erfüllt oder nicht erfüllt. Sie wird aber jetzt gewiß erfüllt sein, da wir ja angenommen haben, daß die x_i^0 eine Lösung des ganzen Systems bilden. — Setzen wir noch $\chi_1(\lambda) \equiv 1$, so gilt $x_i^0 = x_1^0 \chi_i(\lambda)$; $i = 1, 2, \ldots, n$, wobei $\chi_i(\lambda)$ ein nur von $\mathfrak{A}$ abhängiges Polynom genau vom Grade $i-1$ bezeichnet. — Es sei nun λ_j ein Eigenwert, so daß (115) eine Eigenlösung zuläßt. Aus Gründen der Homogenität kann dabei x_1^0 willkürlich gewählt werden. Dadurch sind aber die übrigen x_i^0 eindeutig festgelegt zu $x_i^0 = x_1^0 \chi_i(\lambda_j)$. Es gibt also für $\lambda = \lambda_j$ nur eine Schar von Eigenlösungen, die Anzahl der linear unabhängigen Eigenlösungen ist $= 1$. Da $\mathfrak{A}$ Hermitesch ist, so folgt daraus nach § 13, daß λ_j eine einfache Nullstelle der Säkulargleichung sein muß. Da λ_j ein beliebiger Eigenwert von $\mathfrak{A}$ ist, so sind also alle Eigenwerte einfach. Damit ist gezeigt, daß die Jacobischen Matrizen von einfachem Spektrum sind.

Da der Vektor

$$\mathfrak{x}_k = (\bar{u}_{k1}, \bar{u}_{k2}, \ldots, \bar{u}_{kn}); \qquad \bar{u}_{ki} = x_{1k}^0 \chi_i(\lambda_k), \qquad x_{1k}^0 = \frac{1}{\sqrt{\sum_{\nu=1}^{n} |\chi_\nu(\lambda_k)|^2}}$$

eine normierte Eigenlösung der zu $\lambda_k \mathfrak{E} - \mathfrak{A}$ gehörigen Gleichungen ist und da alle Eigenwerte einfach sind, so ist die Matrix

$$\mathfrak{U} = \| u_{ik} \| = \left\| \frac{\bar{\chi}_k(\lambda_i)}{\sqrt{\sum_{\nu=1}^{n} |\chi_\nu(\lambda_i)|^2}} \right\|$$

nach § 16 unitär und derart, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ eine Diagonalmatrix ist. Die Spektralmatrix von $\mathfrak{A}$ ist daher, wie wir in § 24 gesehen haben,

$$\sigma_{ik}(\mu) = \sum_{\lambda_j \leqq \mu} \bar{u}_{ji} u_{jk}; \qquad i, k = 1, 2, \ldots, n,$$

wofür, da alle Eigenwerte einfach sind, auch $\underset{\lambda_j}{\Delta} \sigma_{ik} = \bar{u}_{ji} u_{jk}$ geschrieben werden kann, so daß

$$\underset{\lambda_j}{\Delta} \sigma_{ik} = \frac{\chi_i(\lambda_j) \bar{\chi}_k(\lambda_j)}{\sum_{\nu=1}^{n} |\chi_\nu(\lambda_j)|^2}$$

ist. Da die (nicht abnehmende) Funktion

$$\varrho(\mu) = \sum_{\lambda_j \leqq \mu} \frac{1}{\sum\limits_{\nu=1}^{n} |\chi_\nu(\lambda_j)|^2}$$

für $\mu = \lambda_j$ den Sprung $\dfrac{1}{\sum\limits_{\nu=1}^{n} |\chi_\nu(\lambda_j)|^2}$ erleidet, so gilt daher

$$(102)' \qquad \mathop{\Delta}_{\mu} \sigma_{ik} = \chi_i(\mu)\, \bar{\chi}_k(\mu) \mathop{\Delta}_{\mu} \varrho$$

für $\mu = \lambda_j$ gewiß; $j = 1, 2, \ldots, n$. Für $\mu \neq \lambda_j$ sind aber beide Funktionen $\sigma_{ik}(\mu)$, $\varrho(\mu)$ stetig, also ist dann $\mathop{\Delta}_{\mu} \sigma_{ik} = \mathop{\Delta}_{\mu} \varrho = 0$, so daß (102)' für alle μ gilt. Damit ist gezeigt, daß jede Jacobische Matrix eine Fouriersche Matrix ist, deren Spektralmatrix eine Darstellung mit orthogonalen Polynomen gestattet. Die Theorie der Hermiteschen Matrizen von einfachem Spektrum ist damit, wie wir in § 31 gesehen haben, zu einem Abschluß gebracht. Wir haben dabei eine sich auf unendliche Matrizen beziehende Untersuchung von Hellinger und Toeplitz (vgl. S. 188 ff.) sinngemäß auf endliche Matrizen umgeschrieben.

§ 33. Fortsetzung. Der assoziierte Kettenbruch und das Momentenproblem.

Die Jacobischen Matrizen hängen mit gewissen Kettenbrüchen aufs engste zusammen. Es sei λ eine Zahl, die außerhalb des Spektrums der Jacobischen Matrix $\mathfrak{A}$ liegt, so daß die zu $\lambda \mathfrak{E} - \mathfrak{A}$ gehörigen inhomogenen Gleichungen

$$(116) \qquad \begin{aligned} (\lambda - a_1) x_1 + b_1 x_2 &= 1, \\ \bar{b}_1 x_1 + (\lambda - a_2) x_2 + b_2 x_2 &= 0, \\ \bar{b}_2 x_2 + (\lambda - a_3) x_3 + b_3 x_4 &= 0, \\ &\cdots \\ \bar{b}_{n-1} x_{n-1} + (\lambda - a_n) x_n &= 0 \end{aligned}$$

gelöst werden können. Die Resolvente $\| R_{ik}(\lambda) \|$ von $\mathfrak{A}$ liefert $x_i = \sum\limits_{k=1}^{n} R_{ik}(\lambda) c_k$ mit $c_1 = 1$; $c_k = 0$, $k > 1$, so daß insbesondere

$$(117) \qquad x_1 = R_{11}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\sigma_{11}(\mu)}{\lambda - \mu} = \int_{-\infty}^{+\infty} \frac{\chi_1(\mu)\, \bar{\chi}_1(\mu)\, d\varrho(\mu)}{\lambda - \mu}$$

gilt. Hierbei ist $\varrho(\mu)$ nicht abnehmend und $\chi_1(\mu)$ eine von Null ver-

schiedene Konstante, etwa C, also $\chi_1(\mu)\bar{\chi}_1(\mu) = |C|^2 > 0$. Wenn wir daher der Bequemlichkeit halber von $\varrho(\mu)$ zu der äquivalenten Dichte $\frac{\varrho(\mu)}{|C|}$ übergehen, so bleibt

$$(118)\qquad x_1 = R_{11}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\varrho(\mu)}{\lambda - \mu} = \sum_{\nu=0}^{\infty} \frac{\mathrm{m}_\nu}{\lambda^{\nu+1}}, \quad \mathrm{m}_\nu = \int_{-\infty}^{+\infty} \mu^\nu \, d\varrho(\mu),$$

wobei die Potenzreihe nur bis zum absolut größten Eigenwerte konvergiert, da

$$(119)\qquad R_{11}(\lambda) = \sum_{j=1}^{n} \frac{\underset{\lambda_j}{\Delta}\varrho}{\lambda - \lambda_j}, \quad \underset{\lambda_j}{\Delta}\varrho > 0$$

gilt. — Andererseits können die Gleichungen (116) in der Gestalt

$$(120)\qquad -x_1 = \frac{1}{a_1 - \lambda - b_1 \frac{x_2}{x_1}}, \quad \frac{x_2}{x_1} = \frac{\bar{b}_1}{a_2 - \lambda - b_2 \frac{x_1}{x_2}}, \ldots,$$

$$\frac{x_{n-1}}{x_{n-2}} = \frac{\bar{b}_{n-2}}{a_{n-1} - \lambda - b_{n-1} \frac{x_n}{x_{n-1}}}, \quad \frac{x_n}{x_{n-1}} = \frac{\bar{b}_{n-1}}{a_n - \lambda}$$

geschrieben werden, sofern wir von den auch sonst höchstens endlich vielen λ-Werten, bei welchen ein Nenner verschwindet, der Kürze halber von vornherein absehen. Aus dem Euklidischen Algorithmus (120) folgt sodann durch sukzessives Einsetzen der Quotienten $\frac{x_{i+1}}{x_i}$ die Kettenbruchentwicklung

$$(121)\qquad x_1 = \cfrac{-\mathrm{m}_0}{a_1 - \lambda - \cfrac{|b_1|^2}{a_2 - \lambda - \cfrac{|b_2|^2}{a_3 - \lambda \; \ddots \; - \cfrac{|b_{n-1}|^2}{a_n - \lambda}}}}; \qquad \mathrm{m}_0 = 1,$$

die also die rationale Funktion (118), (119) darstellt, da (116) nur eine Lösung hat.

Es sei noch ohne Beweis folgendes erwähnt. Damit die unendliche Zahlenfolge $\mathrm{m}_0, \mathrm{m}_1, \ldots$ die Eigenschaft hat, daß eine (und dann auch nur eine) nicht abnehmende Treppenfunktion $\varrho(\mu)$ mit genau n Sprungstellen vorhanden ist derart, daß $\mathrm{m}_\nu = \int_{-\infty}^{+\infty} \mu^\nu \, d\varrho(\mu)$ ausfällt für $\nu = 0, 1, \ldots$ (M o m e n t e n p r o b l e m), ist nicht nur hinreichend, sondern auch notwendig die Existenz einer Jacobischen Matrix, für welche die Relationen (117), (118) bestehen.

Zweites Kapitel.

Analytische Hilfsmittel.

§ 34. Funktionen von beschränkter Schwankung.

Die auf dem endlichen abgeschlossenen Intervall $[a, b]$ erklärte, nicht notwendig reellwertige Funktion $\sigma(\mu)$ heißt daselbst von beschränkter Schwankung, wenn die Summen

$$(122)\quad \sum_{\nu=1}^{N} |\sigma(\xi_{\nu+1}) - \sigma(\xi_\nu)|, \text{ wobei } a = \xi_1 < \cdots < \xi_\nu < \cdots < \xi_{N+1} = b$$

ist, unterhalb einer festen Schranke bleiben. Die kleinste Schranke, d. h. die obere Grenze, wird aus ersichtlichen Gründen (vgl. S. 47) mit $\int\limits_a^b |d\sigma(\mu)|$ bezeichnet; sie nimmt nicht ab, wenn b wächst oder a abnimmt. Ist also $\sigma(\mu)$ auf $[-\infty, +\infty]$ erklärt, so wächst die zu $[a, b]$ gehörige „Totalschwankung" $\int\limits_a^b |d\sigma(\mu)|$ für $a \to -\infty, b \to +\infty$ entweder über alle Grenzen, oder sie strebt einer Zahl zu, die mit $\int\limits_{-\infty}^{+\infty} |d\sigma(\mu)|$ bezeichnet und Totalschwankung von σ auf $[-\infty, +\infty]$ genannt wird; in dem letzteren Falle heißt σ von beschränkter Schwankung auf $[-\infty, +\infty]$. — Ist die Funktion nicht reellwertig, so kann man sie auf genau eine Weise in eine reelle und in eine imaginäre Komponente zerlegen,

$$(123)\qquad \sigma(\mu) = x(\mu) + \sqrt{-1}\, y(\mu).$$

Man überzeugt sich leicht, daß σ dann und nur dann von beschränkter Schwankung ist, wenn x und y es sind. So können wir meist ohne Einschränkung der Allgemeinheit annehmen, daß σ reellwertig ist. — Ist $\sigma(\mu)$ reellwertig und nicht abnehmend[1]), so ist die Summe

[1]) Die Funktion $\sigma(\mu)$ nennen wir nicht abnehmend, wenn sie monoton nicht abnimmt, so daß aus $\mu < \nu$ stets $\sigma(\mu) \leqq \sigma(\nu)$ folgt.

(122) bei jeder Einteilung gleich $\sigma(b) - \sigma(a)$, so daß alsdann die Funktion auf $[-\infty, +\infty]$ dann und nur dann von beschränkter Schwankung ist, wenn sie beschränkt ist; und zwar sind dann die Grenzwerte $\sigma(+\infty)$, $\sigma(-\infty)$ vorhanden und ihre Differenz ist die Totalschwankung.

Es sei $\sigma(\mu)$ reellwertig, von beschränkter Schwankung, jedoch nicht notwendig nicht abnehmend. Es gilt

$$(124)\quad \sigma(\mu) = \varphi(\mu) - \psi(\mu); \quad \varphi(\mu) = \int_{-\infty}^{\mu} |d\sigma(\nu)|, \quad \psi(\mu) = \int_{-\infty}^{\mu} |d\sigma(\nu)| - \sigma(\mu).$$

Man überzeugt sich leicht, daß mit μ nicht nur $\varphi(\mu)$, sondern auch $\psi(\mu)$ nicht abnimmt[1]). Da $\sigma(\mu)$ und $\varphi(\mu)$ von beschränkter Schwankung sind, so ist es auch $\psi(\mu)$. Wir haben damit den C. Jordanschen Satz, wonach die Bedingung, daß die reellwertige Funktion σ von beschränkter Schwankung ist, nicht nur notwendig, sondern auch hinreichend ist dafür, daß σ als Differenz zweier nicht abnehmenden und beschränkten Funktionen dargestellt werden kann. Dies gilt, auch wenn das Intervall endlich ist. — Nach Lebesgue gibt es reellwertige und stetige Funktionen von beschränkter Schwankung, die auf keinem Teilintervall monoton sind (trotzdem sie nach Jordan als Summe zweier monotonen Funktionen dargestellt werden können). — Es gibt stetige Funktionen, die nicht von beschränkter Schwankung sind. Z. B. ist die auf [0, 1] stetige Funktion

$$\sigma_\delta(\mu) = \mu^\delta \sin\frac{1}{\mu}, \qquad 0 < \mu \leqq 1, \qquad \sigma_\delta(0) = 0; \qquad \delta > 0$$

dort dann und nur dann von beschränkter Schwankung, wenn die Reihe $\sum_{\nu=1}^{\infty} \nu^{-\delta}$ konvergiert, und die Weierstraßsche nirgends differentiierbare Funktion ist auf keinem Teilintervall von beschränkter Schwankung. — Da die Grenzwerte $\lim_{\varepsilon \to +0} \sigma(\mu \pm \varepsilon) = \sigma(\mu \pm 0)$ bei jeder monotonen Funktion gewiß vorhanden sind, so sind sie bei jeder Funktion von beschränkter Schwankung vorhanden. Das gleiche gilt für die Grenzwerte $\sigma(+\infty)$, $\sigma(-\infty)$. — Eine Funktion von beschränkter Schwankung kann höchstens abzählbar unendlich viele Unstetigkeitsstellen haben. Denn ist die Funktion nicht abnehmend und von beschränkter Schwankung, so gibt es nur endlich viele Sprungstellen, an welchen der Sprung größer als eine feste Zahl, etwa $> m^{-1}$ ist, weil sonst die Funktion nicht beschränkt sein könnte; setzt man aber $m = 1, 2, \ldots$,

---

[1]) Vgl. z. B. W. Blaschke, Kreis und Kugel, Leipzig, Verlag von Veit & Comp., 1916, S. 26ff.

so durchläuft man alle Sprungstellen. Der Satz gilt also auch für jede Funktion von beschränkter Schwankung, die ja als Differenz zweier monotonen Funktionen dargestellt werden kann. — Offenbar muß die Summe der Sprünge bei jeder Funktion von beschränkter Schwankung absolut konvergieren.

§ 35. Reine Sprungfunktionen. Wesentliche Konvergenz.

Sind $\mu_1, \mu_2, \ldots$ die (eventuellen) Sprungstellen von $\sigma(\mu)$, so ist die Totalschwankung der Funktion

(125) $$\hat{\sigma}(\mu) = \sum_{\mu_j \leqq \mu} \underset{j}{\Delta} \sigma; \qquad \underset{j}{\Delta} \sigma = \sigma(\mu_j + 0) - \sigma(\mu_j - 0)$$

offenbar $\sum_j |\underset{j}{\Delta} \sigma|$, also endlich. Es läßt sich ohne Mühe zeigen, daß die Funktion $\hat{\sigma}(\mu)$ für $\mu \neq \mu_j$ stetig ist und für $\mu = \mu_j$ den Sprung $\underset{j}{\Delta} \sigma$ derart erleidet, und daß $\hat{\sigma}(\mu_j + 0) = \hat{\sigma}(\mu_j)$ ausfällt, d. h. $\hat{\sigma}$ ist auch an den Unstetigkeitsstellen von rechts stetig. Setzt man $\tilde{\sigma}(\mu) = \sigma(\mu) - \hat{\sigma}(\mu)$, also

(126) $$\sigma(\mu) = \tilde{\sigma}(\mu) + \hat{\sigma}(\mu),$$

so ist $\tilde{\sigma}(\mu)$ an den Stetigkeitsstellen von $\sigma(\mu)$ gewiß stetig; es ist ferner

$$\tilde{\sigma}(\mu_j + 0) - \tilde{\sigma}_j(\mu - 0)$$
$$= \sigma(\mu_j + 0) - \sigma(\mu_j - 0) - [\hat{\sigma}(\mu_j + 0) - \hat{\sigma}(\mu_j - 0)] = \underset{j}{\Delta} \sigma - \underset{j}{\Delta} \sigma = 0.$$

Ist nun $\sigma(\mu)$ durchweg von rechts stetig (was z. B. für die Zwecke der Stieltjesschen Integration, wie wir sehen werden, keine Einschränkung der Allgemeinheit bedeutet), so ist

$$\tilde{\sigma}(\mu_j + 0) - \tilde{\sigma}(\mu_j)$$
$$= \sigma(\mu_j + 0) - \sigma(\mu_j) - [\hat{\sigma}(\mu_j + 0) - \hat{\sigma}(\mu_j)] = -[\hat{\sigma}(\mu_j + 0) - \hat{\sigma}(\mu_j)] = 0,$$

da die Sprungfunktion $\hat{\sigma}$ von rechts für alle Fälle stetig ist. Mithin ist $\tilde{\sigma}(\mu)$ überall stetig. Wegen (126) kann also jede von rechts stetige Funktion von beschränkter Schwankung in eine durchweg stetige Funktion von beschränkter Schwankung und in eine „reine Sprungfunktion“ (125), die ebenfalls von beschränkter Schwankung ist, additiv zerlegt werden. Lebesgue und Carathéodory haben die Analyse von σ noch weiter verfolgt, indem sie aus dem stetigen Bestandteil $\tilde{\sigma}$ einen nicht „totalstetigen“ Bestandteil abgespaltet haben. Wir werden darauf in § 50 noch zurückzukommen haben. — Aus (125) geht hervor, daß es Funktionen von beschränkter Schwankung gibt, deren Sprungstellen überall dicht liegen. Man wähle nämlich für $\mu_1, \mu_2, \ldots$ die Folge der

rationalen Zahlen und setze $\underset{j}{\varDelta}\sigma = j^{-2}$. Da aus (125)

$$(127) \qquad \int_{-\infty}^{+\infty} |d\hat{\sigma}(\mu)| = \sum_j |\underset{j}{\varDelta}\sigma|$$

folgt, so ist $\hat{\sigma}(\mu)$ von beschränkter Schwankung. — Da die Sprungstellen μ_j sogar überall dicht liegen können, so kann man sie im allgemeinen nicht der Größe nach ordnen.

Eine Punktmenge der μ-Gerade heißt eine Nullmenge, wenn es zu jedem $\varepsilon > 0$ eine Folge von Intervallen mit der Gesamtlänge ε gibt derart, daß jeder Punkt der Punktmenge in mindestens einem der Intervalle liegt (die auch übereinandergreifen dürfen). Z. B. ist jede (auch überall dichte) abzählbare Menge eine Nullmenge; man wähle nur das j-te Intervall derart, daß es den Punkt μ_j enthält und von der Länge $\frac{\varepsilon}{2^{j+1}}$ ist. Da ein Intervall keine Nullmenge ist, so ist es nicht abzählbar, und das Komplement einer abzählbaren Menge (in bezug auf ein Intervall) ist überall dicht. Die Stetigkeitsstellen einer Funktion von beschränkter Schwankung liegen also überall dicht und „überwiegen" gegenüber den eventuell ebenfalls überall dicht liegenden Unstetigkeitsstellen. — Zwei Funktionen von beschränkter Schwankung, die auf irgendeiner überall dichten Punktmenge übereinstimmen, müssen an allen ihren Stetigkeitsstellen übereinstimmen. Denn zunächst sind die Grenzwerte $\sigma_1(\mu \pm 0)$, $\sigma_2(\mu \pm 0)$ überall vorhanden. Sie können ferner durch $\sigma_\nu(\mu \pm 0) = \lim\limits_{\varepsilon \to +0} \sigma_\nu(\mu \pm \varepsilon)$; $\nu = 1, 2$ erklärt werden, wobei für ε nur solche Werte zugelassen werden, daß $\mu \pm \varepsilon$ in der fraglichen überall dichten Punktmenge liegt. Dann ist aber nach Voraussetzung $\sigma_1(\mu \pm \varepsilon) = \sigma_2(\mu \pm \varepsilon)$. Folglich gilt $\sigma_1(\mu \pm 0) = \sigma_2(\mu \pm 0)$ für alle μ. Ist daher μ insbesondere eine gemeinsame Stetigkeitsstelle, $\sigma_\nu(\mu \pm 0) = \sigma_\nu(\mu)$; $\nu = 1, 2$, so ist, wie behauptet, $\sigma_1(\mu) = \sigma_2(\mu)$. Wir wollen sagen, die beiden Funktionen σ_1, σ_2 von beschränkter Schwankung seien *im wesentlichen identisch* und hierfür $\sigma_1(\mu) [=] \sigma_2(\mu)$ schreiben, wenn $\sigma_1(\mu) = \sigma_2(\mu)$ an allen Stetigkeitsstellen, oder, was nach dem vorhergehenden dasselbe ist, auf einer überall dichten Punktmenge gilt. Hierfür ist offenbar notwendig und hinreichend, daß $\sigma_1(\mu + 0) = \sigma_2(\mu + 0)$ oder $\sigma_1(\mu - 0) = \sigma_2(\mu - 0)$ oder beides für *alle* μ gilt. — Wir wollen sagen, die Funktionenfolge $\sigma^{(1)}(\mu), \sigma^{(2)}(\mu), \ldots$ sei *im wesentlichen konvergent*, wenn es eine Funktion von beschränkter Schwankung, $\sigma(\mu)$, gibt derart, daß $\sigma^{(n)}(\mu) \to \sigma(\mu)$, $n \to +\infty$ an allen Stetigkeitsstellen von $\sigma(\mu)$ gilt; wir schreiben hierfür $\sigma^{(n)}(\mu) [\to] \sigma(\mu)$. Gilt auch $\sigma^{(n)}(\mu) [\to] \sigma^*(\mu)$, so ist $\sigma(\mu) [=] \sigma^*(\mu)$.

Wir wollen den folgenden Satz von Helly beweisen: Ist $\sigma_1(\mu), \sigma_2(\mu), \ldots$ eine Folge von gleichmäßig beschränkten, nicht abnehmenden Funktionen,

$$(128)\quad |\sigma_n(\mu)| < M = \text{konst.};\ \sigma_n(\mu) \leqq \sigma_n(\mu+\varepsilon),\ \varepsilon > 0;\ n = 1, 2, \ldots,$$

so enthält die Folge $\{n\}$ der natürlichen Zahlen eine Teilfolge $\{h_n\}$ derart, daß es eine Funktion $\sigma(\mu)$ von beschränkter Schwankung gibt, gegen welche die Auswahlfolge $\sigma_{h_1}(\mu), \sigma_{h_2}(\mu), \ldots$ im wesentlichen konvergiert: $\sigma_{h_n} [\rightarrow] \sigma$.

§ 36. Der Hellysche Fortpflanzungssatz.

Wir beweisen zuerst den folgenden, eine Art Konvergenzfortpflanzung aussagenden

Hilfssatz. Es sei $\sigma_1(\mu), \sigma_2(\mu), \ldots$ eine Folge von gleichmäßig beschränkten, nicht abnehmenden Funktionen, so daß also (128) gilt, und es möge eine überall dichte Punktfolge P geben derart, daß der Grenzwert $\lim\limits_{n=\infty} \sigma_n(x_p)$ in einem jeden Punkte x_p von P vorhanden ist ($p = 1, 2, \ldots$). Dann konvergiert die Funktionenfolge mindestens im wesentlichen für alle μ, so daß es eine Funktion σ gibt, für welche $\sigma_n [\rightarrow] \sigma$ gilt. Wegen (128) muß σ beschränkt und nicht abnehmend, also von beschränkter Schwankung sein.

Setzen wir

$$(129)\qquad \varrho(\mu) = \lim_{n=+\infty} \sigma_n(\mu);\qquad \mu = x_1, x_2, \ldots, x_p, \ldots.$$

Die Funktion $\varrho(\mu)$ ist also nur auf P erklärt; sie ist auf P wegen (128) nicht abnehmend und beschränkt. Es sei μ eine reelle Zahl, die nicht notwendig auf P liegt. Da P überall dicht ist, so sind in beliebiger Nähe von μ Zahlen $\mu \pm \varepsilon$ vorhanden $(\varepsilon > 0)$, die der Menge P angehören. Da $\varrho(\mu)$ nicht abnimmt, so können wir die Grenzwerte

$$(130)\qquad \sigma(\mu) = \lim_{\varepsilon \to +0} \varrho(\mu+\varepsilon),\qquad \sigma^*(\mu) = \lim_{\varepsilon \to +0} \varrho(\mu-\varepsilon)$$

bilden, wobei also μ beliebig ist und $\mu \pm \varepsilon$ der Menge P angehört. Unser Ziel ist zu beweisen, daß $\sigma_n [\rightarrow] \sigma [=] \sigma^*$ gilt (die Funktionen σ und σ^* sind bereits für alle μ erklärt). Aus (130) und (129) folgt wegen (128)

$$(131)\qquad \sigma^*(\mu) \leqq \sigma(\mu)$$

für alle μ. Wir zeigen, daß hier gewiß das Gleichheitszeichen gilt, wenn an der Stelle μ etwa die Funktion σ (oder σ^*), die offenbar beschränkt und nicht abnehmend ist, stetig ausfällt. Nehmen wir zwecks

Zurückführung ad absurdum an, daß μ eine Stetigkeitsstelle und $\delta = \sigma(\mu) - \sigma^*(\mu)$ dennoch > 0 ist. Dann gibt es zu diesem μ ein $\alpha > 0$ derart, daß

(132) $$\sigma(\nu) - \sigma^*(\mu) \geqq \frac{\delta}{2}$$

ist, sobald $0 < \mu - \nu < \alpha$ gilt. Denn der Stetigkeit zufolge ist $\lim\limits_{\nu\to\mu} \sigma(\nu) = \sigma(\mu)$. Wenn wir nur Werte ν zulassen, die auf der überall dichten, also sich im Punkte μ beiderseits häufenden Menge P liegen, so können wir für (132) wegen (130) auch

(133) $$\lim_{\eta\to+0} \varrho(\mu - \eta) + \frac{\delta}{2} \leqq \lim_{\varepsilon\to+0} \varrho(\nu + \varepsilon), \quad 0 < \mu - \nu < \alpha$$

schreiben. Liegt also $\mu - \eta$ auf P und hinreichend nahe μ, und liegt $\nu + \varepsilon$ auf P und hinreichend nahe an ν, so gilt

(134) $$\varrho(\mu - \eta) + \frac{\delta}{4} \leqq \varrho(\nu + \varepsilon); \quad \varepsilon > 0, \eta > 0,$$

wenn ν auf P und hinreichend nahe an $\mu (> \nu)$ liegt. Es läßt sich also erreichen, daß $\mu - \eta > \nu + \varepsilon$ wird, und trotzdem folgt aus (134)

$$\varrho(\mu - \eta) < \varrho(\nu + \varepsilon).$$

Dies steht im Widerspruch damit, daß die Funktion ϱ (auf P) nicht abnimmt. — Ist also an der Stelle μ eine der beiden Funktionen σ, σ^* stetig, so gilt in (131) das Gleichheitszeichen. Folglich ist $\sigma \,[=]\, \sigma^*$.

Es sei μ beliebig und es seien $\mu + \varepsilon_1$, $\mu - \varepsilon_2$ zwei Punkte von P; $\varepsilon_1 > 0$, $\varepsilon_2 > 0$. Wegen (128), (129), (130) ist

$$\varrho(\mu - \varepsilon_1) = \lim_{n=+\infty} \sigma_n(\mu - \varepsilon_1) \leqq \liminf_{n=+\infty} \sigma_n(\mu) \leqq \limsup_{n=+\infty} \sigma_n(\mu)$$
$$\leqq \lim_{n=+\infty} \sigma_n(\mu + \varepsilon_2) = \varrho(\mu + \varepsilon_2),$$

also

$$\sigma^*(\mu) = \lim_{\varepsilon_1\to+0} \varrho(\mu - \varepsilon_1) \leqq \liminf_{n+\infty} \sigma_n(\mu) \leqq \limsup_{n=+\infty} \sigma_n(\mu)$$
$$\leqq \lim_{\varepsilon_2\to+0} \varrho(\mu + \varepsilon_2) = \sigma(\mu).$$

Ist daher μ eine Stetigkeitsstelle von $\sigma(\mu)$ oder $\sigma^*(\mu)$, also $\sigma(\mu) = \sigma^*(\mu)$, so muß

$$\sigma(\mu) = \liminf_{n=+\infty} \sigma_n(\mu) = \limsup_{n=+\infty} \sigma_n(\mu), \text{ d. h. } \sigma(\mu) = \lim_{n=+\infty} \sigma_n(\mu)$$

sein, womit der von Helly herrührende Hilfssatz (S. 78) bewiesen ist.

§ 37. Das Cantorsche Diagonalprinzip.

Wir erinnern nun an das Cantorsche Diagonalprinzip, das, seitdem seine zentrale Bedeutung insbesondere durch Hilbert erkannt worden ist, zu den wertvollsten Hilfsmitteln der Analysis gehört. Es sei $f_1(x), f_2(x), \ldots$ eine Funktionenfolge, die auf einer abzählbaren Menge M erklärt ist. Es sei x_p ein Punkt von M und es sei die Zahlenfolge $\{f_n(x_p)\}$ bei einem jeden festen p beschränkt. Dann hat diese Zahlenfolge mindestens einen Häufungswert $f(x_p)$; dies besagt, daß die Folge $\{n\}$ der natürlichen Zahlen eine Teilfolge $\{h_n\}$ enthält derart, daß $\lim\limits_{n\to+\infty} f_{h_n}(x_p) = f(x_p)$ gilt. Doch hängt hierbei die Auswahlfolge $\{h_n\}$ von dem Zeiger p des Punktes x_p ab. Das Cantorsche Diagonalprinzip ergibt nun, daß auch eine von p unabhängige Teilfolge $\{h_n\}$ vorhanden ist, so daß $f_{h_n}(x_p) \to f(x_p)$ bei einem jeden p gilt.

Da zunächst die Zahlenfolge $\{f_n(x_1)\}$ beschränkt ist, so enthält sie eine konvergente Teilfolge, die wir mit $\{f_n^{(1)}(x_1)\}$ bezeichnen wollen. Da $\{f_n^{(1)}(x_2)\}$, als eine Teilfolge aus $\{f_n(x_2)\}$, beschränkt ist, so gibt es in $\{f_n^{(1)}(x)\}$ eine Teilfolge $\{f_n^{(2)}(x)\}$, die für $x = x_2$ konvergiert. Sie ist natürlich auch für $x = x_1$ konvergent, da für $x = x_1$ die Folge $\{f_n^{(1)}(x)\}$ konvergent ist, also ist es erst recht ihre Teilfolge. So können wir weitergehen und eine Folge

$$\{f_n^{(1)}(x)\}, \quad \{f_n^{(2)}(x)\}, \ldots, \quad \{f_n^{(m)}(x)\}, \ldots$$

von Folgen konstruieren, welche die folgenden Eigenschaften hat: 1. Die $(m+1)$-te Folge ist eine Auswahlfolge aus der m-ten; 2. die m-te Folge ist für die m Stellen $x = x_1, x_2, \ldots, x_m$ konvergent. — Daraus folgt aber, daß die „Diagonalfolge“ $\{f_n^{(n)}(x)\}$ für *alle* x_m konvergiert. Ist nämlich x_m irgendwie *fest* gewählt, so ist $f_n^{(n)}(x)$ zugleich ein $f_n^{(m)}(x)$, sobald $n \geqq m$ ist. Nun ist aber $\lim\limits_{n=+\infty} f_n^{(m)}(x)$ für die m ersten Stellen x, also auch für x_m vorhanden, also existiert auch $\lim\limits_{n=+\infty} f_n^{(n)}(x)$, da $\{f_n^{(n)}(x)\}$, wenn man von den m ersten $f_{(n)}^{(n)}$ absieht, eine Teilfolge von $\{f_n^{(m)}\}$ ist (m ist fest!). Setzen wir also $f_{h_n} = f_n^{(n)}$, so ist damit die Behauptung des vorigen Absatzes bewiesen.

Es sei nun $\{\sigma_n(\mu)\}$ eine Folge von gleichmäßig beschränkten, nicht abnehmenden Funktionen. Wir wählen auf der μ-Geraden irgendeine abzählbare, jedoch überall dichte Punktmenge P. Da sie abzählbar ist, so gibt es nach dem Diagonalprinzip eine in jedem Punkte von P konvergente Teilfolge $\{\sigma_{h_n}(x)\}$. Da die $\sigma_{h_n}(x)$ offenbar gleichmäßig beschränkt und nicht abnehmend sind, und da P überall dicht liegt,

so ist $\{\sigma_{h_n}\}$ nach dem Hilfssatz im wesentlichen konvergent. Damit ist der bei (128) ausgesprochene Satz bewiesen.

Wir können übrigens etwas weitergehen. Die Folge $\{\sigma_{h_n}(x)\}$ ist höchstens an abzählbar unendlich vielen Stellen, nämlich an den Unstetigkeitsstellen der Grenzfunktion, divergent. Indem man $f_n = \sigma_{h_n}$ setzt, schließt man daraus nach dem Diagonalprinzip, daß $\{\sigma_{h_n}\}$ eine Teilfolge $\{\sigma_{l_n}\}$ enthält, die auch an den Unstetigkeitsstellen konvergiert. Mithin enthält $\{\sigma_n\}$ unter der Voraussetzung (128) eine überall konvergente Teilfolge.

§ 38. Der Hellysche Auswahlsatz und seine funktionentheoretischen Parallelen.

Es sei $\{\sigma_n(\mu)\}$ eine Folge von reellwertigen und gleichmäßig beschränkten Funktionen. Sie brauchen nicht monoton zu sein. Als Ersatz hierfür verlangen wir, daß die Totalschwankung von σ_n unterhalb einer von n unabhängigen Schranke bleibt. Es sei also

$$(135) \qquad |\sigma_n(\mu)| < M, \qquad \int_{-\infty}^{+\infty} |d\sigma_n(\mu)| < K,$$

wobei M und K konstant sind. Zerlegt man σ_n nach Jordan in φ_n und ψ_n, so sind die Funktionen φ_n und ψ_n wegen (124), (135) offenbar gleichmäßig beschränkt. Sie sind ferner, wie wir wissen, nicht abnehmend. Zunächst enthält also $\{\varphi_n\}$ eine überall konvergente Teilfolge $\{\varphi_{l_n}\}$. Da aber die ψ_n, also erst recht die ψ_{l_n}, nicht abnehmend und gleichmäßig beschränkt sind, so enthält $\{\psi_{l_n}\}$ eine überall konvergente Teilfolge $\{\psi_{k_n}\}$. Die Folge $\{\varphi_{k_n}\}$, als Teilfolge von $\{\varphi_{l_n}\}$, ist ebenfalls überall konvergent, also ist es auch die Folge $\{\sigma_{k_n}\} = \{\varphi_{k_n} - \psi_{k_n}\}$. Die Grenzfunktion ist gewiß von beschränkter Schwankung, da die φ_{l_n} und ψ_{l_n} nicht abnehmend und gleichmäßig beschränkt, also von gleichmäßig beschränkter Schwankung sind. — Es war bisher vorausgesetzt, daß die Funktionen $\sigma_n(\mu)$ reellwertig sind. Mit Rücksicht auf (123) folgt jedoch nach einer leicht ersichtlichen Wiederholung der soeben durchgeführten Betrachtung, daß diese Annahme überflüssig ist. Es gilt daher folgender

Auswahlsatz von Helly. Jede, den beiden Bedingungen (135) genügende Funktionenfolge $\{\sigma_n(\mu)\}$ enthält eine im wesentlichen [oder auch überall] konvergente Teilfolge. —

Es sei $\{\sigma_{h_n}\}$ eine im wesentlichen konvergente Teilfolge, $\sigma_{h_n} [\rightarrow] \sigma_*$. Ist $\{\sigma_{l_n}\}$ eine andere, im wesentlichen konvergente Teilfolge aus $\{\sigma_n\}$

mit $\sigma_{l_n}[\rightarrow]\sigma_{**}$, so ist im allgemeinen σ_* nicht $[=]\sigma_{**}$ (so wie eine beschränkte Zahlenfolge zwei verschiedene Häufungswerte haben kann). Gilt bereits $\sigma_n[\rightarrow]\sigma_*$, so daß das Auswahlverfahren überflüssig ist, indem einfach $\sigma_{h_n}=\sigma_n$ gesetzt werden kann, so ist freilich auch $\sigma_{l_n}[\rightarrow]\sigma_*$, also $\sigma_*[=]\sigma_{**}$. — Es gilt nun die folgende Umkehrung: Ergeben zwei im wesentlichen *konvergente* Teilfolgen immer im wesentlichen dieselbe Grenzfunktion, so daß aus $\sigma_{h_n}[\rightarrow]\sigma_*$ und $\sigma_{l_n}[\rightarrow]\sigma_{**}$ bei beliebigen Auswahlfolgen $\sigma_*[=]\sigma_{**}$ folgt, so ist das Auswahlverfahren überflüssig: $\sigma_n[\rightarrow]\sigma_*$.

Den Beweis führen wir nur für den Fall durch, mit dem wir uns später werden zu beschäftigen haben, wenn nämlich alle σ_n reellwertig und nicht abnehmend sind. Es handelt sich also wieder um eine Folge von nicht abnehmenden und gleichmäßig beschränkten Funktionen σ_n. Daß sie im wesentlichen konvergente Teilfolgen enthält, das wissen wir schon. Es wird vorausgesetzt, daß sie die Eigenschaft hat, das aus $\sigma_{h_n}[\rightarrow]\sigma_*$ und $\sigma_{l_n}[\rightarrow]\sigma_{**}$ immer $\sigma_*[=]\sigma_{**}$ folgt, und es wird behauptet, daß dann bereits $\sigma_n[\rightarrow]\sigma_*$ gelten muß.

Zwecks Zurückführung ad absurdum nehmen wir die Unrichtigkeit der Behauptung an. Dann gibt es eine Stetigkeitsstelle μ_0 von σ_* derart, daß $\sigma_n(\mu_0)\rightarrow\sigma_*(\mu_0)$ nicht gilt. Da die Zahlenfolge $\sigma_1(\mu_0)$, $\sigma_2(\mu_0), \ldots$ beschränkt ist, so gibt es daher eine Teilfolge $\{\sigma_{j_n}\}$ mit $\sigma_{j_n}(\mu_0)\rightarrow A\neq\sigma_*(\mu_0)$. Aus $\{\sigma_{j_n}\}$ können wir nach Helly eine im wesentlichen konvergente Teilfolge $\{\sigma_{g_n}\}$ mit einer Grenzfunktion σ_{**} auswählen. Es muß dann $\sigma_{**}(\mu_0)\neq\sigma_*(\mu_0)$ ausfallen. Wegen $\sigma_{h_n}[\rightarrow]\sigma_*$ und $\sigma_{g_n}[\rightarrow]\sigma_{**}$ muß jedoch nach Voraussetzung $\sigma_{**}[=]\sigma_*$ sein. Also gilt $\sigma_{**}(\mu\pm 0)\equiv\sigma_*(\mu\pm 0)$, und zwar sogar an den Unstetigkeitsstellen von σ_* oder σ_{**}. Da μ_0 eine Stetigkeitsstelle von σ_* ist, so muß daher $\sigma_{**}(\mu_0\pm 0)=\sigma_*(\mu_0)$ sein. Nun ist σ_n nicht abnehmend, also ist es auch σ_{**} nicht, folglich gilt $\sigma_{**}(\mu_0-0)\leqq\sigma_{**}(\mu_0)\leqq\sigma_{**}(\mu_0+0)$. Wegen $\sigma_{**}(\mu_0\pm 0)=\sigma_*(\mu_0)$ folgt daraus $\sigma_{**}(\mu_0)=\sigma_*(\mu_0)$, während sich vorher auch $\sigma_{**}(\mu_0)\neq\sigma_*(\mu_0)$ ergeben hat. Unsere Annahmen sind also widerspruchsvoll, w. z. b. w.

Diese Betrachtungen fließen ersichtlich aus zwei verschiedenen Quellen. Die eine ist der Hellysche Hilfssatz auf S. 78, welcher eine Art Konvergenzfortpflanzung ausspricht. Die andere ist das Diagonalprinzip. Dem entspricht es, daß sobald es sich um Funktionenräume handelt, für welche ein Satz über Konvergenzfortpflanzung gilt — wie z. B. bei konvexen Funktionen —, so wird, sobald dabei das Diagonalprinzip im Rechte ist, meistens auch ein Auswahlsatz gelten. — Dem Hellyschen Fortpflanzungssatz und dem Hellyschen Auswahlsatz können wir z. B. in der komplexen Funktionenlehre die folgenden, später zu benutzenden Sätze von Vitali bzw. Montel zur Seite stellen.

Wir wollen unter einem Gebiet Ω der komplexen λ-Ebene ein beschränktes, etwa Jordansches „offenes Gebiet", unter einem Bereich ein Gebiet nebst Berandung verstehen. Der Vitalische Satz, der in einem wichtigen Spezialfall bereits von Stieltjes gefunden worden ist, sagt nun folgendes aus: Sind die Funktionen $f_n(\lambda)$ in Ω regulär und gleichmäßig beschränkt, und ist der Grenzwert $\lim\limits_{n=\infty} f_n(\lambda)$ mindestens auf einer Punktfolge $\lambda_1, \lambda_2, \ldots$ vorhanden, die einen in Ω liegenden Häufungspunkt besitzt, so ist die Funktionenfolge auf jedem, in Ω gelegenen Bereiche konvergent, und zwar gleichmäßig konvergent. Die Grenzfunktion ist also in Ω regulär. — Der Vitalische Fortpflanzungssatz enthält mit Rücksicht auf das Cantorsche Diagonalprinzip offenbar den sogenannten Montelschen Auswahlsatz, wonach jede in Ω reguläre und gleichmäßig beschränkte Funktionenfolge $\{f_n(\lambda)\}$ eine in einem beliebigen, in Ω gelegenen Bereiche gleichmäßig konvergente Teilfolge $\{f_{h_n}(\lambda)\}$ enthält. Das Auswahlverfahren ist wieder dann und nur dann überflüssig, d. h. man darf dann und nur dann $h_n = n$ setzen, wenn jede in Ω konvergente Teilfolge eine und dieselbe Grenzfunktion besitzt.

§ 39. Der Stieltjessche Integralbegriff.

Es sei $f(\mu)$ eine in dem endlichen abgeschlossenen Intervall $a \leqq \mu \leqq b$, kürzer $[a, b]$, erklärte stetige Funktion, und $\sigma(\mu)$ eine Funktion von beschränkter Schwankung; für $f(x)$ braucht dies nicht zuzutreffen, und $\sigma(\mu)$ darf Sprünge haben. Wir behaupten, daß dann $f(\mu)$ in bezug $\sigma(\mu)$ in dem auf S. 47 erklärten Sinne gewiß integrabel ist. Man bilde nämlich, wie auf S. 47, eine Stieltjessche Näherungssumme

$$S = \sum_{q=1}^{N+1} f(\nu_{q-1})\,[\sigma(\mu_q) - \sigma(\mu_{q-1})]; \tag{136}$$

$$a = \mu_0 < \mu_1 < \cdots < \mu_{N+1} = b, \quad \mu_{q-1} \leqq \nu_{q-1} \leqq \mu_q,$$

und eine andere,

$$S' = \sum_{q=1}^{N'+1} f(\nu'_{q-1})\,[\sigma(\mu'_q) - \sigma(\mu'_{q-1})]; \tag{137}$$

$$a = \mu'_0 < \mu'_1 < \cdots < \mu'_{N'+1} = b, \quad \mu'_{q-1} \leqq \nu'_{q-1} \leqq \mu'_q,$$

und beachte, daß in Anbetracht der gleichmäßigen Stetigkeit von $f(x)$ zu jedem $\varepsilon^* > 0$ ein $\eta > 0$ gefunden werden kann derart, daß die Ungleichheit $|\mu^{(1)} - \mu^{(2)}| \leqq \eta$ die weitere Ungleichheit $|f(\mu^{(1)}) - f(\mu^{(2)})| \leqq \varepsilon^*$

[1]) Wegen der Durchführung der Einzelheiten vgl. z. B. die Darstellung von W. Blaschke, l. c. (s. S. 75).

zur Folge hat. In den Elementen wird daraus geschlossen, daß im Spezialfall $\sigma(\mu) \equiv \mu$ der gewöhnlichen Integration $|S - S'| \leqq \varepsilon^* (b - a)$ gilt, sobald jede der Differenzen $\mu_q - \mu_{q-1}$, $\mu'_q - \mu'_{q-1}$ hinreichend klein ist. Hierbei ist $b - a$ die Totalschwankung von $\sigma(\mu) \equiv \mu$. Man erkennt aus dieser Bemerkung leicht, daß die wörtliche Wiederholung jener Schlußweise für unseren Fall folgendes ergibt: Es gibt zu jedem $\varepsilon > 0$ ein $\eta > 0$ derart, daß aus $\mu_q - \mu_{q-1} < \eta$ und $\mu'_q - \mu'_{q-1} < \eta$

$$|S - S'| \leqq \varepsilon = \varepsilon^* \int_a^b |d\sigma(\mu)|$$

folgt. Dies besagt aber, nach Definition (S. 47), die Existenz des Grenzwertes $\lim S = \int_a^b f(\mu)\, d\sigma(\mu)$. — Analog läßt sich beweisen, daß die Näherungssummen $\sum_{q=1}^{N+1} |\sigma(\mu_q) - \sigma(\mu_{q-1})|$ der Totalschwankung in demselben Sinne der Totalschwankung $\int_a^b |d\sigma(\mu)|$ zustreben. Daraus folgt unmittelbar, daß $\int_a^x |d\sigma(\mu)|$, als Funktion von x betrachtet, an jeder Stetigkeitsstelle x von $\sigma(\mu)$ stetig ist, während das Integral beim Passieren einer Unstetigkeitsstelle x von $\sigma(\mu)$ den Sprung $|\sigma(x + 0) - \sigma(x - 0)| > 0$ erleidet, so daß zusammenfassend

$$\underset{x}{\Delta} \int_a^x |d\sigma(\mu)| = |\underset{x}{\Delta}\sigma| \tag{138}$$

gilt, im Einklang mit (127). Entsprechend gilt in leicht verständlicher Bezeichnungsweise

$$\begin{aligned} \int_a^{a+0} f(\mu)\, d\sigma(\mu) &= f(a)[\sigma(a+0) - \sigma(a)], \\ \int_{b-0}^b f(\mu)\, d\sigma(\mu) &= f(b)[\sigma(b) - \sigma(b-0)], \end{aligned} \tag{139}$$

also

$$\underset{x}{\Delta} \int_a^x f(\mu)\, d\sigma(\mu) = f(x) \underset{x}{\Delta}\sigma = \left(\int_{x-0}^x + \int_x^{x+0}\right) f(\mu)\, d\sigma(\mu), \tag{140}$$

ferner

$$\int_a^b = \int_a^{a+0} + \int_{a+0}^{b-0} + \int_{b-0}^b, \tag{141}$$

wie es sich aus (63), S. 48, durch leicht ersichtliche Grenzübergänge ergibt. Ist $\hat{\sigma}(\mu)$ eine reine Sprungfunktion (S. 76), deren sämtliche Sprungstellen μ_j im Innern von $[a, b]$ liegen, so ist

$$(142) \qquad \int_a^b f(\mu)\, d\hat{\sigma}(\mu) = \sum_{j=1}^{\infty} f(\mu_j) \mathop{\Delta}_{\mu_j} \hat{\sigma},$$

wobei $\sum |\mathop{\Delta}_{\mu_j} \hat{\sigma}|$, wie wir wissen, konvergiert. Also ist auch die Reihe (142) absolut konvergent (denn f ist stetig, also beschränkt). Der Vorbehalt, daß in (142) alle Sprungstellen im Innern von $[a, b]$ liegen, ist wesentlich; denn Sprünge am Rande müssen wegen (136), (140), (141) auf eine besondere Weise berücksichtigt werden. — Bezeichnet M das Maximum von $|f(\mu)|$ auf $[a, b]$, so ist

$$(143) \qquad \left|\int_a^b f(\mu)\, d\sigma(\mu)\right| \leqq M \int_a^b |d\sigma(\mu)|.$$

Dies folgt daraus, daß die entsprechende Ungleichheit bereits für die Näherungssummen S richtig ist. Ist σ auf $[-\infty, +\infty]$ von beschränkter Schwankung, $\int_{-\infty}^{+\infty} |d\sigma(\mu)| < +\infty$, so folgt aus (143) unmittelbar, daß $\int_{-\infty}^{+\infty} f(\mu)\, d\sigma(\mu)$ in dem auf S. 47 erklärten Sinne gewiß existiert, wenn f (außer auf jedem $[a, b]$ stetig zu sein) auf $[-\infty, +\infty]$ beschränkt ist. Doch ist dies keine notwendige Bedingung. Wir kommen darauf gelegentlich noch zurück (vgl. S. 102). — Ist auf $[-\infty, +\infty]$ die Funktion $f(\mu)$ stetig und beschränkt und $\sigma(\mu)$ von beschränkter Schwankung (also gewiß beschränkt), so können die vorher entwickelten Formeln durch den Grenzübergang $a \to -\infty$, $b \to +\infty$ unmittelbar auf den Fall unendlicher Integrationsgrenzen übertragen werden. — Es sei noch die Ungleichheit

$$(144) \qquad \left|\int_a^b f(\mu)\, g(\mu)\, d\sigma(\mu)\right|^2 \leqq \int_a^b |f(\mu)|^2 d\sigma(\mu) \cdot \int_a^b |g(\mu)|^2 d\sigma(\mu)$$

erwähnt, wobei f, g stetig, σ nicht abnehmend ist. Diese Abschätzung folgt aus der Schwarzschen Ungleichung, wenn man den Grenzübergang von den Näherungssummen S zu den Integralen vornimmt.

§ 40. Fortsetzung. Partielle Integration. Unempfindlichkeit der inneren Sprungstellen.

Man hat, wenn $\nu_0 = \mu_0 = a$ und $\nu_N = \mu_{N+1} = b$ gewählt wird,

$$S = \sum_{q=1}^{N+1} f(\nu_{q-1})[\sigma(\mu_q) - \sigma(\mu_{q-1})] = \sum_{q=1}^{N+1} f(\nu_{q-1})\,\sigma(\mu_q) - \sum_{q=0}^{N} f(\nu_q)\,\sigma(\mu_q)$$

$$= f(b)\,\sigma(b) - f(a)\,\sigma(a) - \sum_{q=1}^{N} \sigma(\mu_q)[f(\nu_q) - f(\nu_{q-1})]$$

(Abelsche Summation), wobei die letzte Summe [wegen $a = \nu_0 \leqq \mu_1 \leqq \cdots \leqq \nu_{q-1} \leqq \mu_q \leqq \nu_q \leqq \cdots \leqq \nu_N = b$ und da $\mu_{q+1} - \mu_q > 0$, also $\nu_{q+1} - \nu_q > 0$ gilt] eine zu $\int\limits_a^b \sigma(\mu)\, d f(\mu)$ gehörige Stieltjessche Näherungssumme darstellt. Folglich ist auch dieses Integral vorhanden, da $\lim S$ existiert. Und zwar hat man

$$\text{(145)}\qquad \int\limits_a^b f(\mu)\, d\sigma(\mu) = f(b)\,\sigma(b) - f(a)\,\sigma(a) - \int\limits_a^b \sigma(\mu)\, d f(\mu).$$

Vorausgesetzt ist dabei, daß f stetig und σ von beschränkter Schwankung ist. Mit $f(\mu) \equiv \mu$ erhält man daraus mit Rücksicht auf eine Bemerkung auf S. 47, daß die Funktionen von beschränkter Schwankung im Riemannschen Sinne integrierbar sind, was freilich wohlbekannt ist. Die Formel (145) ist diejenige der partiellen Integration. Geht man zur Grenze $a \to -\infty$, $b \to +\infty$ über und nimmt man an, daß σ auch auf $[-\infty, +\infty]$ von beschränkter Schwankung und f daselbst beschränkt ist, so ist, wie wir wissen, $\int\limits_{-\infty}^{+\infty} f\, d\sigma$ vorhanden, und (145) geht über in

$$\text{(145)'}\qquad \int\limits_{-\infty}^{+\infty} f(\mu)\, d\sigma(\mu) = -\int\limits_{-\infty}^{+\infty} \sigma(\mu)\, d f(\mu),$$

wenn z. B. $f(\mu) \to 0$, $\mu \to \pm\infty$ vorausgesetzt wird. Doch genügt bereits z.B. die Annahme $f(\mu) \to 0$ für $\mu \to +\infty$, da es bei der Stieltjesschen Integration auf eine additive Konstante der Belegung nicht ankommt, so daß $\sigma(-\infty) = 0$, d. h. $\sigma(a) \to 0$, $a \to -\infty$ angenommen werden kann, während f beschränkt ist, usf.

Ändert man den Wert der Funktion $\sigma(\mu)$ an Stellen, die im Innern von $[a, b]$ liegen, derart ab, daß die abgeänderte Funktion, $\sigma^*(\mu)$, ebenfalls von beschränkter Schwankung ist und mit σ auf einer überall

dicht liegenden Punktmenge übereinstimmt, so wird dadurch der Wert des Integrals $\int\limits_a^b f\,d\sigma$ für kein f geändert. Denn das Integral $\int\limits_{a+0}^{b-0}$ kann — da es eben existiert — unter ausschließlicher Benutzung von Teilsummen berechnet werden, deren Einteilungspunkte jener auf $a < x < b$ liegenden, überall dichten Punktmenge angehören, und andererseits gelten die Formeln (139), (141). — Die Endpunkte nehmen freilich wegen (141) eine Sonderstellung ein. Im Falle $a = -\infty$, $b = +\infty$ hat man jedoch keine Endpunkte. Für alle Fälle ist $\sigma(a) = \sigma^*(a)$, $\sigma(b) = \sigma^*(b)$, $\sigma(\mu)\,[=]\,\sigma^*(\mu)$ hinreichend für $\int\limits_a^b f\,d\sigma = \int\limits_a^b f\,d\sigma^*$. An endlich vielen, im Innern von $[a, b]$ liegenden Stellen kann σ immer beliebig abgeändert werden. Hat σ unendlich viele Sprungstellen $\mu_1, \mu_2, \ldots$, so sind die Reihen $\sum\limits_j |\sigma(\mu_j \pm 0) - \sigma(\mu_j)|$, $\sum\limits_j |\sigma(\mu_j + 0) - \sigma(\mu_j - 0)|$ gewiß konvergent, da sonst σ nicht von beschränkter Schwankung wäre. Setzt man

$$\sigma^*(\mu) = \sigma(\mu) + \sum_{\mu_j \leqq \mu} (\Theta_j \sigma(\mu_j + 0) - \Theta_j \sigma(\mu_j - 0)) + \cdots,$$

so ist daher, wie wir wissen (S. 77), σ^* von beschränkter Schwankung und $[=]\,\sigma$, sofern die Θ beschränkt sind. Indem man sie passend gleich $+1$, 0, oder -1 wählt, kann nach den Bemerkungen auf S. 76 erreicht werden, daß etwa $\sigma^*(\mu) = \sigma^*(\mu + 0)$ auch an den Unstetigkeitsstellen von σ gilt. Für die Zwecke der Stieltjesschen Integration (und ebenso bei der Berechnung der Totalschwankung $\int\limits_a^b |d\sigma|$) kann daher immer angenommen werden, daß σ mit Ausnahme der eventuellen, im Endlichen liegenden Randpunkte durchweg von rechts stetig ist.

Wir können jetzt aus (145) die folgende Folgerung ziehen: Ist f stetig, σ von beschränkter Schwankung und in den Endpunkten von $[a, b]$ stetig, so gilt

$$(146)\quad \int\limits_a^b f(\mu)\,d\sigma(\mu) = \int\limits_a^b f(\mu)\,d\int\limits_a^\mu d\sigma(\nu) \equiv \int\limits_a^b f(\mu)\,d\varrho(\mu),\ \varrho(\mu) = \int\limits_a^\mu d\sigma(\nu).$$

Denn zunächst gilt wegen (145) oder mit Rücksicht auf die Definition des Stieltjesschen Integrals [S. 47; vgl. S. 84]

$$(146)'\qquad \sigma(\mu) = \int\limits_a^\mu d\sigma(\nu) \equiv \varrho(\mu)$$

an den Stetigkeitsstellen von $\sigma(\mu)$ gewiß. Ersetzt man also σ in (145) an allen vier Stellen durch ϱ, so erhält man eine Gleichung, deren rechte Seite mit der rechten Seite der unter (145) stehenden Gleichung übereinstimmt. Also müssen auch die linken Seiten identisch sein, womit (146) bewiesen ist.

Hat σ eine Ableitung σ', so geht die Näherungssumme S nach dem Mittelwertsatz der Differentialrechnung in $\sum_{q=1}^{N+1} f(\nu_{q-1})\,\sigma'(\lambda_{q-1})\,(\mu_q - \mu_{q-1})$ über, wobei λ_{q-1} eine im offenen Intervall $\mu_{q-1} < \lambda_{q-1} < \mu_q$ gelegene Zahl bedeutet; da ν_q nur an die Bedingung $\mu_{q-1} \leqq \nu_{q-1} \leqq \mu_q$ gebunden ist (weil das Integral $\int_a^b f\,d\sigma$ eben existiert), so können wir $\nu_{q-1} = \lambda_{q-1}$ wählen, so daß die S Riemannsche Näherungssummen von $f(\mu)\,\sigma'(\mu)$ werden. Ist also außer f auch σ' stetig, so gilt, da dann $f\sigma'$ gewiß im Riemannschen Sinne integrierbar ist,

$$\int_a^b f(\mu)\,d\sigma(\mu) = \int_a^b f(\mu)\,\sigma'(\mu)\,d\mu. \tag{147}$$

Doch wird σ im allgemeinen Sprünge haben, und auch wenn σ durchweg stetig ist, braucht σ' nicht zu existieren; ferner kann (147), auch wenn σ' existiert, jedoch nicht stetig ist, unrichtig sein. Es ist also keine Rede davon, daß das Stieltjessche Integral immer auf ein Riemannsches zurückgeführt werden kann.

§ 41. Der Hellysche Satz über gliedweise Integration.

Es seien f; $f_1, f_2, \ldots$ stetige Funktionen auf $[a, b]$ und es sei σ von beschränkter Schwankung. Daraus, daß $f_n \to f$ bei einem jeden μ gilt, folgt

$$\int_a^b f_n(\mu)\,d\sigma(\mu) \to \int_a^b f(\mu)\,d\sigma(\mu) \tag{148}$$

bekanntlich nicht einmal im Falle $\sigma(\mu) \equiv \mu$, während das gleichmäßige Bestehen von $f_n \to f$ für

$$\int_a^b f_n(\mu)\,d\mu \to \int_a^b f(\mu)\,d\mu \tag{149}$$

hinreichend, wenn auch nicht notwendig ist. Analog liegen die Verhältnisse bei (148). Wir begnügen uns mit dem einfachsten hinreichenden Kriterium: gilt $f_n \to f$ gleichmäßig, so ist (148) gewiß richtig. Dann

gibt es nämlich zu jedem $\varepsilon > 0$ ein n_0 derart, daß für $n > n_0$ die Differenz $f_n(\mu) - f(\mu)$ absolut $\leqq \frac{\varepsilon}{\int_a^b |d\sigma(\mu)|}$ ist, also, wegen (143),

$$\left|\int_a^b f_n(\mu)\,d\sigma(\mu) - \int_a^b f(\mu)\,d\sigma(\mu)\right| = \left|\int_a^b [f_n(\mu) - f(\mu)]\,d\sigma(\mu)\right| \leqq \varepsilon$$

gilt (der Nenner kann nur im trivialen Fall $\sigma = \text{const.}$ verschwinden).

Unvergleichbar tiefer liegt ein Satz von Helly, der die Frage behandelt, wann die Grenzgleichung

$$\int_a^b f(\mu)\,d\sigma_n(\mu) \to \int_a^b f(\mu)\,d\sigma(\mu) \tag{150}$$

gewiß richtig ist; hierbei sind also die Belegungen nicht fest, während f eine feste stetige Funktion bezeichnet. Der Hellysche Satz besagt, daß (150) sicher gilt, wenn die Bedingungen

$$\begin{gathered} \sigma_n [\to] \sigma, \quad \sigma_n(a) \to \sigma(a), \quad \sigma_n(b) \to \sigma(b), \\ \int_a^b |d\sigma_n(\mu)| < K, \quad |\sigma_n(\mu)| < K \end{gathered} \tag{151}$$

erfüllt sind, wobei K eine beliebige feste Schranke bezeichnet. Das Wesentliche ist dabei, daß das gleichmäßige Bestehen von $\sigma_n(\mu) \to \sigma(\mu)$ nicht vorausgesetzt wird; eine derartige einschränkende Bedingung wäre auch an den Hellyschen Auswahlsatz, wo keine gleichmäßige Konvergenz behauptet werden kann, gar nicht angepaßt.

Wir nehmen zunächst an, es sei bereits bewiesen, daß (150) eine Folge von (151) ist, falls f stetig differentiierbar ist, und zeigen, daß dann der Satz allgemein, d. h. für jede stetige Funktion f, gelten muß. Da nämlich f stetig ist, so gibt es eine Folge $\{f_m(\mu)\}$ von stetig differentiierbaren Funktionen, die gleichmäßig gegen f konvergieren, denn nach Weierstraß kann man f sogar durch Polynome gleichmäßig approximieren. Es gibt also zu jedem $\varepsilon > 0$ ein m_0 derart, daß $|f - f_m| \leqq \varepsilon$ für alle μ gilt, sobald $m \geqq m_0$ gewählt wird. Mithin ist wegen (143)

$$\left|\int_a^b f\,d\sigma - \int_a^b f_m\,d\sigma\right| \leqq \varepsilon \int_a^b |d\sigma| \quad \text{für} \quad m \geqq m_0,$$

und ebenso gilt wegen (143) und (151)

$$\left|\int_a^b f_m\,d\sigma_n - \int_a^b f\,d\sigma_n\right| \leqq \varepsilon \int_a^b |d\sigma_n| \leqq \varepsilon K \quad \text{für} \quad m \geqq m_0,$$

wie auch dabei n beschaffen sein mag. Endlich ist (150) nach Voraussetzung für jede stetig differentiierbare Funktion f richtig, sofern (151) gilt. Wir können daher zu jedem m ein n_m derart bestimmen, daß

$$\left|\int_a^b f_m \, d\sigma_n - \int_a^b f_m \, d\sigma\right| \leqq \varepsilon \quad \text{für} \quad n \geqq n_m$$

gewiß gilt. Setzen wir in diesen drei Abschätzungen $m = m_0$ und $N = n_{m_0}$ und addieren sie, so folgt

$$\left|\int_a^b f \, d\sigma - \int_a^b f \, d\sigma_n\right| \leqq \varepsilon \left(\int_a^b |d\sigma| + K + 1\right) \quad \text{für} \quad n \geqq N.$$

Da hierbei der Faktor von ε unabhängig von n ist, so ist damit gezeigt, daß der Satz (150) unter der Voraussetzung (151) allgemein richtig ist, wenn er für stetig differentiierbare Funktionen f richtig ist.

Man erinnere sich nun an den folgenden, im wesentlichen von Arzelà herrührenden Satz: Ist $\{f_n(x)\}$ eine auf $[a, b]$ erklärte, gleichmäßig beschränkte Folge von integrierbaren Funktionen, und gibt es eine beschränkte und integrierbare Funktion $f(x)$ derart, daß $f_n(x) \to f(x)$ an allen Stetigkeitsstellen von $f(x)$ oder eventuell mit Ausnahme einer Nullmenge gilt, so ist (149) gewiß richtig; die Integration ist in dem alten, Riemannschen Sinne zu verstehen. — Doch gehört der Satz bereits zu denjenigen, die der Natur der Sache nach erst im Rahmen der Lebesgueschen Theorie eine Abrundung erfahren konnten: der Satz erreicht — vom *heutigen* Standpunkt aus[1]) — ein künstlich eingeschränktes Ziel mit einem gekünstelt primitiven und entsprechend unbequemen Apparat. Es hat also keinen Sinn, auf den Beweis dieses Teilergebnisses einzugehen.

Ist nun $\{\sigma_n(\mu)\}$ eine Funktionenfolge derart, daß die Bedingungen (151) erfüllt sind, und ist $f'(\mu)$ eine stetige Funktion, so sind die Funktionen $\sigma_n(\mu) f'(\mu)$ integrierbar und gleichmäßig beschränkt. Die Grenzfunktion der σ_n, die Funktion $\sigma(\mu)$, ist von beschränkter Schwankung, also ist auch $\sigma(\mu) f'(\mu)$ integrabel und es gilt wegen $\sigma_n [\to] \sigma$ an allen Stetigkeitsstellen der Grenzfunktion offenbar auch $\sigma_n f' \to \sigma f'$, so daß nach dem Satze von Arzelà

$$\int_a^b \sigma_n(\mu) f'(\mu) \, d\mu \to \int_a^b \sigma(\mu) f'(\mu) \, d\mu,$$

---

[1]) Die von Arzelà (1885) benutzten Intervallsätze gehören freilich zu den wichtigsten Ergebnissen des vorigen Jahrhunderts auf diesem Gebiet.

also wegen (147)

$$\int_a^b \sigma_n(\mu)\, df(\mu) \rightarrow \int_a^b \sigma(\mu)\, df(\mu)$$

gilt; da wegen (151) auch die Grenzgleichungen

$$\sigma_n(a) f(a) \rightarrow \sigma(a) f(a) \quad \text{und} \quad \sigma_n(b) f(b) \rightarrow \sigma(b) f(b)$$

bestehen, so muß mit Rücksicht auf die Formel (145) der partiellen Integration auch (150) gelten, w. z. b. w.

§ 42. Die Resolventenintegrale.

Es sei $\sigma(\mu)$ eine auf $[-\infty, +\infty]$ erklärte Funktion, die reellwertig und auf $[-\infty, +\infty]$ von beschränkter Schwankung ist, und es sei λ eine nicht reelle Zahl. Die reelle Komponente von λ möge mit $\mathfrak{R}[\lambda]$, die imaginäre mit $\mathfrak{J}[\lambda]$ bezeichnet werden, so daß $\lambda = \mathfrak{R}[\lambda] + \sqrt{-1}\,\mathfrak{J}[\lambda]$ ist. Die Funktion $(\lambda - \mu)^{-1}$ ist auf $[-\infty, +\infty]$ stetig und beschränkt. Mit Rücksicht auf eine oben (S. 86) gemachte Bemerkung ist daher das Integral

$$R(\lambda) = \int_{-\infty}^{+\infty} \frac{d\sigma(\mu)}{\lambda - \mu}, \tag{152}$$

das wir aus später ersichtlichen Gründen als das zu σ gehörige Resolventenintegral bezeichnen wollen, gewiß vorhanden. Wegen (145)′ kann man hierfür auch

$$R(\lambda) = -\int_{-\infty}^{+\infty} \frac{\sigma(\mu)\, d\mu}{(\mu - \lambda)^2} \tag{153}$$

schreiben. Wir zeigen, daß die durch diese Formel dargestellte Funktion für nicht reelle λ gewiß regulär ist. Es bezeichne zu diesem Zwecke K einen Kreisbereich der λ-Ebene, der mit der Achse des Reellen keinen Punkt gemein hat, so daß es ein $d > 0$ gibt derart, daß $|\lambda - \mu| \geqq d$ für alle λ auf K und für alle $\mu \left(\gtrless 0\right)$ gilt. Man bilde bei festem a und b die zu dem Stieltjesschen Integral

$$T_a^b(\lambda) = \int_a^b \frac{d\sigma(\mu)}{\lambda - \mu}$$

gehörigen Näherungssummen $S = S(\lambda)$. Sie konvergieren für alle λ von K gleichmäßig gegen das Integral $T_a^b(\lambda)$. Denn die Funktion

$f(\mu) = (\lambda - \mu)^{-1}$ ist für $a \leqq \mu \leqq b$ auf K gleichmäßig stetig, so daß das ε auf S. 84 für alle λ gleichmäßig gewählt werden kann. Da $S(\lambda)$ eine rationale Funktion ist, deren sämtliche Pole auf der reellen Achse liegen, so ist sie auf K gewiß regulär, so daß im Innern von K auch die Grenzfunktion $T_a^b(\lambda)$ regulär ist (denn die Grenzfunktion einer in einem Gebiete gleichmäßig konvergenten Folge von Funktionen, die in einem Gebiet regulär sind, ist in diesem Gebiete gewiß regulär).

Nun gilt aber die definitorische Grenzgleichung $T_a^b(\lambda) \to R(\lambda)$; $a \to -\infty$, $b \to +\infty$ ebenfalls gleichmäßig auf K. Denn in der aus (143) folgenden Abschätzung

$$\left| \int_b^{+\infty} \frac{d\sigma(\mu)}{\lambda - \mu} \right| \leqq \frac{1}{d} \int_b^{+\infty} | d\sigma(\mu) | \tag{154}$$

konvergiert die Schranke für $b \to +\infty$ wegen $\int\limits_{-\infty}^{+\infty} | d\sigma(\mu) | < +\infty$ gegenNull, und die Schranke ist unabhängig von λ. Da entsprechendes auch für $\int\limits_{-\infty}^{a}$ gilt, so ist $R(\lambda)$ für alle nicht reellen λ regulär.

Ist die reelle Zahl μ_0 derart, daß $\sigma(\mu)$ in einer hinreichend kleinen Umgebung $\mu_0 - \varepsilon < \mu < \mu_0 + \varepsilon$ von μ_0 konstant, $\equiv \sigma(\mu_0)$ ist, so ist $R(\lambda)$ auch für $\lambda = \mu_0$ regulär. Denn zunächst kann dann das Integral durch

$$\left(\int_{-\infty}^{\mu_0 - \varepsilon - 0} + \int_{\mu_0 + \varepsilon + 0}^{+\infty} \right) \frac{d\sigma(\mu)}{\lambda - \mu} \tag{155}$$

ersetzt werden, da eine Konstanzstrecke der Belegung zu einem Stieltjesschen Integral nichts beiträgt. Da alle in einem hinreichend kleinen, um den Punkt μ_0 geschlagenen Kreise liegenden λ derart sind, daß $| \lambda - \mu | >$ konst. > 0 gilt für alle μ, die in die Integrale (155) eingehen, so kann die frühere Betrachtung offenbar wiederholt werden.

Sind λ_1 und λ_2 zwei Punkte, deren Verbindungsstrecke durch die Achse des Reellen nicht geschnitten wird, so gilt

$$\int_{\lambda_1}^{\lambda_2} d\lambda \int_{-\infty}^{+\infty} \frac{d\sigma(\mu)}{\lambda - \mu} = \int_{-\infty}^{+\infty} d\sigma(\mu) \int_{\lambda_1}^{\lambda_2} \frac{d\lambda}{\lambda - \mu}, \tag{156}$$

die Integrationen geradlinig verstanden. Denn zunächst tragen die Umgebungen der beiden Punkte $\mu = \pm + \infty$ zu den beiden zwei-

fachen Integralen wieder beliebig wenig bei (vgl. die vorige Schlußweise). Es kommt also nur darauf an, die Richtigkeit der Formel

$$\int_{\lambda_1}^{\lambda_2} d\lambda \int_a^b \frac{d\sigma(\mu)}{\lambda-\mu} = \int_a^b d\sigma(\mu) \int_{\lambda_1}^{\lambda_2} \frac{d\lambda}{\lambda-\mu}$$

einzusehen. Hier können aber wegen (147) beide Stieltjesschen Integrale in Riemannsche verwandelt werden, und die Vertauschbarkeit der Reihenfolge der beiden Integrationen ist für diesen Fall aus den Elementen bekannt.

Ist $\sigma(\mu)$ eine Treppenfunktion mit den Sprungstellen $\lambda_1, \lambda_2, \ldots, \lambda_n$ und mit den Sprüngen $\underset{1}{\Delta}\sigma, \underset{2}{\Delta}\sigma, \ldots, \underset{n}{\Delta}\sigma$, so ist (152) eine rationale gebrochene Funktion mit einfachen Polen,

$$R(\lambda) = \sum_{\nu=1}^{n} \frac{\underset{\nu}{\Delta}\sigma}{\lambda-\lambda_\nu},$$

wobei die Funktion σ durch $R(\lambda)$ an allen ihren Stetigkeitsstellen bis auf eine additive Konstante eindeutig bestimmt wird. Dies besagt ja nur, daß mit $R(\lambda)$ auch die Pole und die Residuen von $R(\lambda)$ bekannt sind. Wir wollen nun eine Stieltjessche Formel herleiten, welche für den allgemeinen Fall entsprechendes leistet. Es wird sich dabei eine eigentümliche Verallgemeinerung der Residuenformel

$$\lim_{\lambda \to \lambda_\nu} (\lambda - \lambda_\nu) R(\lambda) = \underset{\nu}{\Delta}\sigma$$

ergeben. Doch handelt es sich nunmehr nicht um die Sprünge allein; denn σ kann jetzt durchweg stetig und nirgends konstant sein. — Wir werden sodann nach Hilbert auch die allgemeine „Residuenformel" herleiten.

§ 43. Die Stieltjessche Umkehrformel.

Da $\sigma(\mu)$ für reellwertig vorausgesetzt ist, nimmt die Funktion (152), (153) für konjugiert komplexe Argumente konjugiert komplexe Werte an, $R(\bar{\lambda}) = \bar{R}(\lambda)$, so daß wir uns ohne Einschränkung der Allgemeinheit auf die obere Halbebene $\mathfrak{J}[\lambda] > 0$ beschränken können. Setzen wir noch zur Abkürzung

$$p = \mathfrak{R}[\lambda], \quad q = \mathfrak{J}[\lambda],$$

so ist wegen

$$R(\lambda) = \int_{-\infty}^{+\infty} \frac{d\sigma(\mu)}{\lambda-\mu} = \int_{-\infty}^{+\infty} \frac{(\bar{\lambda}-\mu)\, d\sigma(\mu)}{|\lambda-\mu|^2};$$

$$|\lambda-\mu|^2 = (p-\mu)^2 + q^2, \quad \bar{\lambda} = p - \sqrt{-1}\, q$$

der imaginäre Teil der Zahl $R(\lambda)$ offenbar

$$\mathfrak{J}[R(\lambda)] = -q \int_{-\infty}^{+\infty} \frac{d\sigma(\mu)}{(p-\mu)^2+q^2}.$$

Wir integrieren diesen Ausdruck bei festem $q = \mathfrak{J}[\lambda] > 0$ in bezug auf die von q unabhängig gedachte Variable $p = \mathfrak{R}[\lambda]$ etwa von 0 bis $\alpha \gtrless 0$, und beachten, daß dabei die Vertauschung der sich auf μ und p beziehenden Integrationen gewiß gestattet ist. Man erkennt dies ebenso wie die Richtigkeit der Formel (156). Setzt man daher

$$F(\alpha, q) = F(\alpha, \mathfrak{J}[\lambda]) = \int_0^\alpha \mathfrak{J}[R(\lambda)]\, d\mathfrak{R}[\lambda], \tag{157}$$

so gilt

$$F(\alpha, q) = \int_{-\infty}^{+\infty} d\sigma(\mu) \int_0^\alpha \frac{-q\,dp}{(p-\mu)^2+q^2},$$

wobei

$$\int_0^\alpha \frac{dp}{(p-\mu)^2+q^2} = \frac{1}{q}\operatorname{arc\,tg}\frac{\alpha-\mu}{q} - \frac{1}{q}\operatorname{arc\,tg}\frac{-\mu}{q} \qquad \left(-\frac{\pi}{2} < \operatorname{arc\,tg} x < \frac{\pi}{2}\right)$$

ist. Hieraus folgt, wenn

$$G(\alpha, q) = \int_{-\infty}^{+\infty} \operatorname{arc\,tg}\frac{\alpha-\mu}{q}\, d\sigma(\mu) \tag{158}$$

gesetzt wird, offenbar

$$F(\alpha, q) = G(0, q) - G(\alpha, q). \tag{159}$$

Wegen (145) ist aber

$$\begin{aligned} &\int_{-a}^{+a} \operatorname{arc\,tg}\frac{\alpha-\mu}{q}\, d\sigma(\mu) \\ &= \sigma(a)\operatorname{arc\,tg}\frac{\alpha-a}{q} - \sigma(-a)\operatorname{arc\,tg}\frac{\alpha+a}{q} + \int_{-a}^{+a} \frac{1}{q}\,\frac{\sigma(\mu)\,d\mu}{1+\left(\frac{\alpha-\mu}{q}\right)^2}, \end{aligned} \tag{160}$$

wobei das Integral rechts, da α und q feste Zahlen sind und $\sigma(\mu)$ auf $[-\infty, +\infty]$ beschränkt ist, für $a \to +\infty$ der Grenze

$$H(\alpha, q) = \int_{-\infty}^{+\infty} \frac{1}{q}\,\frac{\sigma(\mu)\,d\mu}{1+\left(\frac{\alpha-\mu}{q}\right)^2} \tag{161}$$

zustrebt. Da σ von beschränkter Schwankung ist, so sind die beiden Grenzwerte $\sigma(+\infty)$, $\sigma(-\infty)$ vorhanden. Wegen $q = \mathfrak{J}[\lambda] > 0$ gilt endlich $\operatorname{arctg}\frac{\alpha \mp a}{q} \to \mp \frac{\pi}{2}$ für $a \to +\infty$, so daß (160) für $a \to +\infty$ wegen (158) und (161) in

$$G(\alpha, q) = -\sigma(+\infty)\frac{\pi}{2} - \sigma(-\infty)\frac{\pi}{2} + H(\alpha, q),$$

und daher (159) in

$$F(\alpha, q) = H(0, q) - H(\alpha, q) \tag{162}$$

übergeht.

Wir zeigen nun, daß

$$\lim_{q \to +0} \frac{2}{\pi} H(\alpha, q) = \sigma(\alpha + 0) + \sigma(\alpha - 0) \tag{163}$$

gilt. Mit Rücksicht auf (161) beweisen wir mehr, wenn wir zeigen, daß

$$\sigma(\alpha+0) = \lim_{q \to +0} \frac{2}{\pi} \int_{\alpha}^{+\infty} \frac{1}{q} \frac{\sigma(\mu)\, d\mu}{1 + \left(\frac{\alpha-\mu}{q}\right)^2}, \quad \sigma(\alpha-0) = \lim_{q \to +0} \frac{2}{\pi} \int_{-\infty}^{\alpha} \frac{1}{q} \frac{\sigma(\mu)\, d\mu}{1 + \left(\frac{\alpha-\mu}{q}\right)^2} \tag{164}$$

gilt, wobei wir uns auf die erste dieser beiden Formeln beschränken können. Sie kann in der Gestalt

$$\sigma(\alpha + 0) = \lim_{q \to +0} \frac{2}{\pi} \int_{0}^{+\infty} \frac{1}{q} \frac{\sigma(\mu + \alpha)\, d\mu}{1 + \frac{\mu^2}{q^2}} \tag{165}$$

geschrieben werden und geht, wenn man

$$\tau(\mu) = \sigma(\mu + \alpha) - \sigma(\alpha + 0)$$

setzt und

$$\frac{2}{\pi} \int_{0}^{+\infty} \frac{1}{q} \frac{d\mu}{1 + \frac{\mu^2}{q^2}} = \frac{2}{\pi} \operatorname{arctg} x \Big|_{0}^{+\infty} = 1 \qquad \left(x = \frac{\mu}{q}\right) \tag{166}$$

beachtet, in

$$0 = \lim_{q \to +0} \int_{0}^{+\infty} \frac{1}{q} \frac{\tau(\mu)\, d\mu}{1 + \frac{\mu^2}{q^2}} \tag{167}$$

über, wobei $\tau(\mu)$ eine auf $[0, +\infty]$ beschränkte und für $\mu \to +0$ der Bedingung $\tau(\mu) \to 0$ Genüge leistende Funktion bezeichnet [den

Umstand, daß die Funktion τ nicht nur beschränkt und derart ist, daß die Grenzwerte $\tau(\pm 0)$ existieren, sondern daß sie außerdem von beschränkter Schwankung ist, brauchen wir beim Beweis von (167) nicht heranzuziehen]. Es gibt also zu jedem $\varepsilon > 0$ ein $\delta > 0$ derart, daß $|\tau(\mu)|$ für $0 \leqq \mu \leqq \delta$ kleiner als ε ist, und es gibt eine Schranke M derart, daß $|\tau(\mu)| \leqq M$ auch für $\mu \geqq \delta$ gilt. Beachtet man daher, daß in (167) der Faktor von τ, der sogenannte Kern, durchweg > 0 ist, so folgt einerseits

$$\left|\int_0^\delta \frac{1}{q} \frac{\tau(\mu)\, d\mu}{1+\frac{\mu^2}{q^2}}\right| < \varepsilon \int_0^\delta \frac{1}{q} \frac{d\mu}{1+\frac{\mu^2}{q^2}} < \frac{\varepsilon}{2},$$

wenn δ hinreichend klein gewählt wird [vgl. (166)], und andererseits

$$\left|\int_\delta^{+\infty} \frac{1}{q} \frac{\tau(\mu)\, d\mu}{1+\frac{\mu^2}{q^2}}\right| < M \int_\delta^{+\infty} \frac{q\, d\mu}{q^2+\mu^2} < q M \int_\delta^{+\infty} \frac{d\mu}{\mu^2} < \frac{\varepsilon}{2},$$

wenn nach erfolgter Wahl von δ auch q hinreichend klein gewählt wird. Durch Addition der beiden letzten Formeln folgt (167). Damit ist also (163) bewiesen. Diese Schlußweise ist zuerst von Hamilton und Dirichlet, später (bei positiven Kernen) von Schwarz und Weierstraß verwendet worden und ist seitdem sehr geläufig. — Dieser Zusammenhang mit der Theorie der „singulären Integrale“ geht aus den üblichen Beweisen der Umkehrformel nicht so klar hervor.

Aus (157), (161), (162), (163) folgt nun, wenn noch μ an Stelle von α geschrieben wird, die angekündigte Stieltjessche Umkehrformel

(168)
$$-\frac{2}{\pi} \lim_{\mathfrak{J}[\lambda] \to +0} \int_0^\mu \mathfrak{J}[R(\lambda)]\, d\mathfrak{R}[\lambda]$$
$$= \sigma(\mu+0) + \sigma(\mu-0) - [\sigma(+0) + \sigma(-0)],$$

wobei die linke Seite nur von μ abhängt. Die Belegung $\sigma(\mu)$ von (152) ist also an allen ihren Stetigkeitsstellen bis auf eine der Natur der Sache nach unbestimmt bleibende additive Konstante $[\sigma(+0)+\sigma(-0)]$ vollständig bestimmt. Legen wir die additive Konstante etwa durch die Normierung $\sigma(-\infty) = 0$ fest, so sind daher alle Funktionen $\sigma(\mu)$, die vermittels (152) dieselbe analytische Funktion $R(\lambda)$ erklären, im wesentlichen identisch, d. h. einander $[=]$. Es gibt unter diesen Funktionen eine und nur eine, die von rechts durchweg stetig ist.

§ 44. Die Hilbertsche „Residuenformel".

Zu der Stieltjesschen Umkehrformel analog ist die folgende, die von Hilbert herrührt:

$$(169)\quad \lim_{\varepsilon \to +0} \sqrt{-1}\,\varepsilon\, R(\mu + \varepsilon \sqrt{-1}) = \underset{\mu}{\Delta}\sigma = \sigma(\mu+0) - \sigma(\mu-0);\quad \varepsilon > 0.$$

Dies ist eben die auf S. 93 erwähnte „Residuenformel". Es ist jedoch zu beachten, daß in dieser Formel nur senkrechte Annäherung an die reelle Achse in Betracht kommt, was durch die Kompliziertheit der im Punkte $\lambda = \mu$ möglichen Singularität von $R(\lambda)$ bedingt ist. Die Funktion $R(\lambda)$ wird wegen (169) bei senkrechter Annäherung an die reelle Achse genau von der ersten Ordnung unendlich, wenn μ eine Sprungstelle der Belegung σ ist, an den Stetigkeitsstellen von σ von einer niedrigeren Ordnung oder [wie an den Konstanzstellen von σ, wo $R(\lambda)$ nach S. 92 regulär ist] überhaupt nicht. Dies entspricht dem Umstand, daß für eine Treppenfunktion σ, d. h. für ein rationales R, alle Pole von R einfach sind.

Als Sprungstellen von σ bezeichnen wir nur die Stellen μ mit $\underset{\mu}{\Delta}\sigma \neq 0$. „Hebbare" Unstetigkeiten, in denen $\sigma(\mu+0) = \sigma(\mu-0) \neq \sigma(\mu)$ ist, sind ja für die Stieltjessche Integration nach § 87 belanglos.

Wir bemerken zunächst, daß es genügt, die Richtigkeit von (169) für den Fall nachzuweisen, wo μ eine Stetigkeitsstelle von σ ist, unter Stetigkeitsstelle wieder eine Stelle μ mit $\underset{\mu}{\Delta}\sigma = 0$ verstanden. Es sei nämlich μ_0 eine Sprungstelle von σ. Es bezeichne σ' diejenige Treppenfunktion, die an der Stelle μ_0 den Sprung $\underset{\mu_0}{\Delta}\sigma$ erleidet und sonst stetig ist, und man setze $\sigma'' = \sigma - \sigma'$. Die zu σ' und σ'' gehörigen Resolventenintegrale mögen R' bzw. R'' heißen, so daß $R = R' + R''$ gilt. Dann ist σ'' an der Stelle μ_0 mit Rücksicht auf die Definition von σ' gewiß stetig. Also ist (169) für R'', σ'' nach Voraussetzung richtig, mithin gilt $\lim \varepsilon R''(\mu_0 + \sqrt{-1}\,\varepsilon) = 0$ für $\varepsilon \to +0$. Es genügt also zu zeigen, daß $\lim \sqrt{-1}\,\varepsilon R'(\mu_0 + \sqrt{-1}\,\varepsilon) = \underset{\mu_0}{\Delta}\sigma$ gilt. Diese Behauptung ist aber trivial; denn nach der Definition von σ' ist

$$R'(\lambda) = \frac{\underset{\mu_0}{\Delta}\sigma}{\lambda - \mu_0}.$$

Wir können also beim Beweis von (169) annehmen, daß μ eine Stetigkeitsstelle von σ ist. Wir können ferner durch eine Verlegung des Nullpunktes der μ-Geraden erreichen, daß $\mu = 0$ wird.

Man kann daher $\mu = 0$ und $\underset{0}{\Delta}\sigma = \underset{\mu}{\Delta}\sigma = 0$ setzen, wodurch (169) in $\varepsilon R(\varepsilon\sqrt{-1}) \to 0$ übergeht. Die Voraussetzung ist dann, daß die Belegung den beiden Bedingungen $\sigma(\pm\varepsilon) \to 0$ genügt. Wegen (153) ist aber

$$|\varepsilon R(\varepsilon\sqrt{-1})| \leqq \varepsilon \int_{-\infty}^{+\infty} \frac{|\sigma(\mu)|\,d\mu}{|\mu - \varepsilon\sqrt{-1}|^2} = \int_{-\infty}^{+\infty} \frac{1}{\varepsilon} \frac{|\sigma(\mu)|\,d\mu}{1 + \frac{\mu^2}{\varepsilon^2}},$$

und Integral rechts konvergiert wegen (164) und mit Rücksicht auf $\sigma(\pm\varepsilon) \to 0$ für $\varepsilon \to +0$ gegen Null. Damit ist (169) bewiesen.

Geht man in (168) von λ zu $\bar\lambda$ über, so ändert sich dadurch $\mathfrak{R}[\lambda]$ nicht, während $R(\lambda)$ in $\bar R(\lambda)$, also $\mathfrak{J}[R(\lambda)]$ in $-\mathfrak{J}[R(\lambda)]$ übergeht, so daß $2\mathfrak{J}[R(\lambda)] = \mathfrak{J}[R(\lambda) - R(\bar\lambda)]$ ist. Schreibt man noch $\int_\alpha^\mu$ an Stelle von $\int_0^\mu$, was nur auf eine Verlegung des Nullpunktes hinausläuft, so folgt

$$(170) \quad \begin{aligned} &-\frac{1}{\pi} \lim_{\mathfrak{J}[\lambda]\to+0} \int_\alpha^\mu \mathfrak{J}[R(\lambda) - R(\bar\lambda)]\, d\mathfrak{R}[\lambda] \\ &= \sigma(\mu+0) + \sigma(\mu-0) - [\sigma(\alpha+0) + \sigma(\alpha-0)]. \end{aligned}$$

Wir erwähnen noch, daß diese Formeln von Hellinger auf den Matrizenkalkül übertragen worden sind, wo an Stelle von $R(\lambda)$ eine Matrix von solchen „Resolventenintegralen" tritt.

§ 45. Funktionentheoretisches über die Resolventenintegrale.

Ist α eine Konstanzstelle der Belegung, so daß es ein $\delta > 0$ gibt derart, daß für $\alpha - \delta < \mu < \alpha + \delta$ identisch $\sigma(\mu) = \sigma(\alpha)$ gilt, so ist, wie wir auf S. 92 gesehen haben, *die durch die Formel* (152) *dargestellte Funktion* $R(\lambda)$ in dem um den Punkt α geschlagenen Kreise vom Radius δ regulär. Aber auch nur dann. Denn ist sie in diesem Kreise regulär, so ist sie in jedem noch kleineren Kreise vom Radius δ' gleichmäßig stetig, so daß dabei $R(\lambda) - R(\bar\lambda)$ für $\lambda - \bar\lambda \to 0$, d. h. für $\mathfrak{J}[\lambda] \to 0$ gleichmäßig gegen Null konvergiert. Aus (170) folgt aber dann $\sigma(\mu+0) + \sigma(\mu-0) \equiv \sigma(\alpha+0) + \sigma(\alpha-0)$, also, wenn man selbstverständlich von hebbaren Unstetigkeiten absieht, offenbar auch $\sigma(\mu) \equiv \sigma(\alpha)$ für $|\mu - \alpha| < \delta'$, w. z. b. w. Dieser Satz rührt von Grommer her. Danach sind die Nichtkonstanzstellen der Belegung $\sigma(\mu)$ mit den im Endlichen gelegenen Singularitäten der *durch die Formel* (152)

dargestellten Funktion $R(\lambda)$ identisch, indem alle nicht reellen λ gewiß reguläre Stellen sind (S. 92).

Wir sprechen dabei von der Funktion, die durch die Formel (152) dargestellt wird, und nicht von der analytischen Funktion, die aus dieser Funktion, wobei etwa $\mathfrak{J}[\lambda] > 0$ sein möge, durch analytische Fortsetzung entsteht. Denn es ist z.B. möglich, daß die Formel (152) für $\mathfrak{J}[\lambda] > 0$ und für $\mathfrak{J}[\lambda] < 0$ zwei *verschiedene* analytische Funktionen darstellt, so daß also $R(\lambda) - R(\bar{\lambda}) \to 0$ für $\lambda - \bar{\lambda} \to 0$ nicht gilt, trotzdem beide Funktionen ganze transzendente Funktionen sind, also im Endlichen überhaupt keine Singularitäten haben. Nur wird dann die eine Funktion durch die Formel (152) nur in der oberen, die andere nur in der unteren Halbebene dargestellt. Ein Beispiel ist das zur nirgends konstanten Belegung

$$\sigma(\mu) = \int_{-\infty}^{\mu} e^{-t^2}\, dt$$

gehörige Integral $R(\lambda)$. Es ist freilich auch der andere Grenzfall möglich, daß die nirgends konstante Belegung sich so verhält, daß die reelle Achse für die beiden [wegen $R(\bar{\lambda}) = \bar{R}(\lambda)$ einfach: konjugiert komplexen] Funktionen eine natürliche Grenze bildet. Zwischen diesen beiden Fällen steht der dritte, wo $\sigma(\mu)$ etwa für $\mu < -1$ und auch für $\mu > 1$ konstant, also

$$R(\lambda) = \int_{-1-0}^{1+0} \frac{d\sigma(\mu)}{\lambda - \mu} \tag{152'}$$

ist. Dann können die beiden Funktionen, nämlich $R(\lambda)$ und $R(\bar{\lambda})$, ineinander gewiß analytisch fortgesetzt werden. Denn jede reelle Zahl, die absolut > 1 ist, ist eine reguläre Stelle der durch (152)′ dargestellten Funktion. Alle Punkte des Intervalles $-1 \leqq \mu \leqq 1$ können dabei wesentlich singuläre Stellen sein, so daß eine Fortsetzung von $R(\lambda)$ in $R(\bar{\lambda})$ durch einen *unmittelbaren* Übergang durch die Punkte des Intervalles $-1 \leqq \mu \leqq 1$ unmöglich ist. Doch können alle Punkte des Intervalls $-1 < \mu < 1$ auch reguläre Stellen sein, trotzdem dort σ nirgends konstant ist. Nur müssen dann die beiden Punkte $\mu = \pm 1$ Verzweigungspunkte sein, so daß $R(\lambda)$ bei unmittelbarer Überschreitung eines Punktes des Intervalles $-1 < \mu < 1$ nicht den Wert $R(\bar{\lambda})$ erhält, der sich bei der Umkreisung des Intervalles $-1 \leqq \mu \leqq 1$ ergibt, sondern einen andern Zweig der mehrdeutigen analytischen Funktion liefert, so daß also $-1 \leqq \mu \leqq 1$ mit Hinblick auf den Grommerschen Satz als ein Verzweigungsschlitz aufzufassen ist. Das einfachste

Beispiel ist die Gaußsche Kettenbrucherzeugende der Legendreschen Polynome,

$$R(\lambda) = \int_{-1}^{1} \frac{d\mu}{\lambda - \mu} = \log \frac{\lambda + 1}{\lambda - 1};$$

$$\sigma(\mu) \equiv -1 \quad \text{für} \quad \mu \leqq -1; \qquad \sigma(\mu) \equiv \mu \quad \text{für} \quad -1 \leqq \mu \leqq +1;$$
$$\sigma(\mu) \equiv 1 \quad \text{für} \quad \mu \geqq 1.$$

Dies ist eben das Beispiel, durch das Gauß, wie aus seinem berühmten Briefe an Bessel (1811) hervorgeht, viele Jahrzehnte vor Cauchy bzw. Puiseux auf die einschlägigen Grundtatsachen der Funktionentheorie geführt wurde.

Abschließende Ergebnisse über die sich im Anschluß an (152) erhebenden funktionentheoretischen Fragen scheinen in der Literatur nicht vorzuliegen. Es würde sich vor allem um die Frage handeln, auf welche Weise die Art der Singularitäten von $R(\lambda)$ durch das infinitesimale Verhalten der Belegung bestimmt wird und wie überhaupt eine vorgegebene Funktion $R(\lambda)$ beschaffen sein muß, damit ein Zweig von ihr eine Darstellung (152) mit einer passend zu wählenden Belegung gestattet. — Bereits diejenigen $R(\lambda)$, die eine Darstellung mit einer reinen Sprungfunktion [vgl. (142)] gestatten, so daß

$$(171) \qquad R(\lambda) = \sum_{\nu=1}^{\infty} \frac{\Delta_\nu \sigma}{\lambda - \lambda_\nu}; \qquad \sum_{\nu=1}^{\infty} |\Delta_\nu \sigma| < +\infty$$

gilt, sind, wenn die Sprungstellen etwa überall dicht liegen, nach Goursat und Pringsheim wesentlich komplizierter, als man es von vornherein anzunehmen geneigt wäre.

Ist λ reell und gleich μ_0, wobei μ_0 keine Konstanzstelle von σ ist, so hat das Integral (152) zunächst keinen Sinn; als Cauchyscher Hauptwert z. B. kann es jedoch existieren. Wir können hier auf diese Fragen, die aus leicht ersichtlichen Gründen (vgl. S. 96) wieder in die Theorie der Integraldarstellung willkürlicher Funktionen hinüberführen, nicht näher eingehen. — Es sei nur erwähnt, daß die Stieltjessche Umkehrformel mit derjenigen von Fourier durch eine Art Laplacesche Transformation zusammenhängt.

§ 46. Fortsetzung. Das Verhalten im Unendlichen. Momente und C. Neumannsche Reihen.

Eine eventuelle matrizentheoretische Bedeutung der durch die Formel (152) nicht dargestellten Zweige ist nicht bekannt. Wir wollen fortan unter $R(\lambda)$ die durch (152) *unmittelbar dargestellte* Funktion ver-

stehen, wodurch wir sie künstlich eindeutig machen. Die singulären Stellen dieser Zweige sind also gewiß reell, und sie sind, wenn die Punkte der einen unmittelbaren Übergang verhindernden Schlitze zu den singulären Stellen hinzugezählt werden, mit den Nichtkonstanzstellen der Belegung identisch. Allerdings haben wir noch den Beweis nachzuholen, daß in diesem Satz der unendlich ferne Punkt (der evtl. aufgeschlitzten und so für alle Fälle schlichten) λ-Ebene keine Sonderstellung einnimmt. Liegt nicht die gesamte Belegung im Endlichen, gibt es also keine Schranke $\varrho > 0$ derart, daß für $\mu \leqq -\varrho$ identisch $\sigma(\mu) = \sigma(-\infty)$ und für $\mu \geqq \varrho$ ebenso $\sigma(\mu) = \sigma(+\infty)$ gilt, so häufen sich die Nichtkonstanzstellen von σ, also die Singularitäten von R, bei $\lambda = \infty$, so daß dann $\lambda = \infty$ gewiß singulär ist. Wir haben also nur die Umkehrung zu beweisen: Liegt die ganze Belegung ganz im Endlichen, so daß bei hinreichend großem ϱ

$$(172) \qquad R(\lambda) = \int_{-\varrho}^{\varrho} \frac{d\sigma(\mu)}{\lambda - \mu}$$

gilt, so ist $\lambda = \infty$ eine reguläre Stelle der Funktion $R(\lambda)$. Der Beweis führt auf die C. Neumannschen Reihen bzw. auf die bereits früher betrachteten *Momente*

$$(173) \qquad \mathfrak{m}_n = \int_{-\infty}^{+\infty} \mu^n \, d\sigma(\mu); \qquad n = 0, 1, 2, \ldots .$$

Wegen (173) und da σ für $\mu \leqq -\varrho$ und $\mu \geqq \varrho$ konstant ist, gilt $|\mathfrak{m}_n| = |\int_{-\varrho}^{\varrho} \mu^n \, d\sigma(\mu)| \leqq \varrho^n \int_{-\varrho}^{\varrho} |d\sigma(\mu)|$, so daß die Reihe $\sum \mathfrak{m}_n \lambda^{-n-1}$ für $|\lambda| > \varrho$ konvergiert. Wir haben nur noch zu zeigen, daß sie die Funktion $R(\lambda)$ darstellt. Nun ist aber die Reihe $\sum \mu^n \lambda^{-n-1}$ für $-\varrho \leqq \mu \leqq \varrho$ (bei jedem λ mit $|\lambda| > \varrho$) gleichmäßig konvergent, also ist nach (148) die gliedweise Integration

$$\int_{-\varrho}^{+\varrho} \sum_{n=0}^{\infty} \frac{\mu^n}{\lambda^{n+1}} \, d\sigma(\mu) = \sum_{n=0}^{\infty} \int_{-\varrho}^{+\varrho} \frac{\mu^n}{\lambda^{n+1}} \, d\sigma(\mu) = \sum_{n=0}^{\infty} \frac{\mathfrak{m}_n}{\lambda^{n+1}}$$

gewiß gestattet, und linkerhand steht offenbar die Funktion (172).

Um die Vorstellungen zu fixieren, nehmen wir jetzt an, daß σ nicht abnehmend, $d\sigma = |d\sigma|$ ist. Von Momenten spricht man meistens nur in diesem Fall. Gibt es keine Schranke ϱ, so kann, wie das Beispiel $\sigma(\mu) = \mu^3$ zeigt, bereits das Integral $\mathfrak{m}_1$ divergieren. Doch kann, wie die auf S. 99 erwähnte exponentielle Belegung zeigt, jedes Integral $\mathfrak{m}_n$ konvergent sein, ohne daß die ganze Belegung im Endlichen

liegt, d. h. ohne daß $R(\lambda)$ die für hinreichend große $|\lambda|$ *konvergente* Entwicklung $\sum\limits_0^\infty \mathfrak{m}_n \lambda^{-n-1}$ zuläßt. Hingegen stellt diese C. Neumannsche Reihe (sofern sie überhaupt gebildet werden kann, d. h. wenn alle Integrale $\mathfrak{m}_n$ existieren) die Funktion $R(\lambda)$ in einem asymptotischen Sinne des Wortes immer dar. Ist nämlich λ nicht reell, so ist das Integral (152) konvergent und wegen

$$\frac{1}{\lambda-\mu} = \sum_{\nu=0}^{n-1} \frac{\mu^\nu}{\lambda^{\nu+1}} + \frac{1}{\lambda^n}\frac{\mu^n}{\lambda-\mu}$$

gilt

$$R(\lambda) - \sum_{\nu=0}^{n-1} \frac{\mathfrak{m}_\nu}{\lambda^{\nu+1}} = \frac{1}{\lambda^n}\int_{-\infty}^{+\infty} \frac{\mu^n \, d\sigma(\mu)}{\lambda-\mu}, \tag{174}$$

wobei letzteres Integral, wegen $|d\sigma| = d\sigma$ und da alle $\mathfrak{m}_n$ vorhanden sind, ersichtlich gegen Null strebt, wenn die Variable λ derart ins Unendliche geht, daß sie außerhalb eines die reelle Achse in seinem Innern enthaltenden Parallelstreifens bleibt. Denn wegen (173) und nach der Schwarzschen Ungleichung (144) ist $\int\limits_{-\infty}^{+\infty} |\mu^n| \, d\sigma(\mu) \leqq \sqrt{\mathfrak{m}_0 \mathfrak{m}_{2n}}$, also endlich; das Restglied, nämlich die rechte Seite von (174), konvergiert daher bei diesem Grenzübergang stärker nach Null als λ^{-n}, also stärker, als die in (174) linkerhand berücksichtigten Terme, und dies gilt bei einem jeden n. Mit dem Poincaréschen Zeichen der Asymptotizität gilt also $R(\lambda) \sim \sum\limits_{\nu=0}^{\infty} \frac{\mathfrak{m}_\nu}{\lambda^{\nu+1}}$, wenn λ auf die angegebene Weise unendlich wird. — Setzt man der Einfachheit halber voraus, daß σ eine stetige Ableitung σ' hat, und bezeichnet man $d_n = |\sqrt[n]{\mathfrak{m}_n}|$ als das n-te *reduzierte* Moment von σ, so erkennt man leicht, daß das reduzierte Moment mit n um so „stärker" wächst, je „langsamer" die Funktion $|\sigma'(\mu)|$ für $\mu \to \pm\infty$ ihrem Limes inferior, der $=0$ ist, zustrebt. Setzt man z. B. $\sigma'(\mu) = e^{-|\mu|^\delta}$, $\delta > 0$, so strebt $\frac{\sqrt[2n]{d_{2n}}}{n^\delta}$ einer von 0 und $+\infty$ verschiedenen Grenze zu, wenn $n \to +\infty$.

§ 47. Der Grommersche Konvergenzsatz.

Es sei $\{\sigma_n(\mu)\}$ eine Folge von nicht abnehmenden, gleichmäßig beschränkten Funktionen auf $[-\infty, +\infty]$, die im wesentlichen gegen $\sigma(\mu)$ konvergieren. Bezeichnet M die obere Grenze aller

$\int\limits_{-\infty}^{+\infty} |d\sigma_n(\mu)|$, so ist, wenn etwa $\sigma_n(-\infty) = 0$ angenommen wird,

$$(175) \qquad |\sigma_n(\mu)| \leqq M, \quad d\sigma_n \geqq 0, \quad \int\limits_{-\infty}^{+\infty} |d\sigma_n(\mu)| \leqq M; \quad \sigma_n [\rightarrow] \sigma.$$

Es bezeichne $R_n(\lambda)$ das zur Belegung σ_n gehörige Resolventenintegral, und $R(\lambda)$ das zur Grenzbelegung gehörige. Wir beweisen nach Grommer mit einer von Hamburger herrührenden Ergänzung, daß dann für alle nicht reellen λ die Funktion $R_n(\lambda)$ gegen das zur Grenzbelegung gehörige Resolventenintegral konvergiert: $R_n(\lambda) \rightarrow R(\lambda)$, $n \rightarrow +\infty$.

Zunächst ist wegen (153)

$$|R_n(\lambda) - R(\lambda)| = \left| \int\limits_{-\infty}^{+\infty} \frac{\sigma_n(\mu) - \sigma(\mu)}{(\lambda - \mu)^2} d\mu \right| \leqq \int\limits_{-\infty}^{+\infty} \frac{|\sigma_n(\mu) - \sigma(\mu)|}{|\lambda - \mu|^2} d\mu.$$

Bezeichnet also p den Realteil, q den Imaginärteil von λ, so gilt

$$(176) \qquad |R_n(\lambda) - R(\lambda)| \leqq \left(\int\limits_{-\infty}^{a} + \int\limits_{a}^{b} + \int\limits_{b}^{+\infty} \right) \frac{|\sigma_n(\mu) - \sigma(\mu)|}{(p - \mu)^2 + q^2} d\mu,$$

wobei wir die Zerlegung von $[-\infty, +\infty]$ durch a und b in drei Stücke auf die folgende Weise festlegen wollen. Es sei $\varepsilon > 0$ beliebig klein. Da p und q feste reelle Zahlen sind und $q \neq 0$ ist, so können wir a so nahe an $-\infty$ wählen, daß

$$\int\limits_{-\infty}^{a} \frac{d\mu}{(p - \mu)^2 + q^2} \leqq \frac{\varepsilon}{6M}$$

wird. Wegen (175) ist nicht nur $|\sigma_n|$, sondern auch $|\sigma|$ gewiß $\leqq M$, also $|\sigma_n - \sigma| \leqq 2M$ für alle μ, so daß das erste Integral (176) offenbar $\leqq \frac{\varepsilon}{6M} 2M = \frac{\varepsilon}{3}$ ist, und zwar für alle n. Ebenso können wir durch geeignete Wahl von b erreichen, daß das Integral

$$\int\limits_{b}^{+\infty} \leqq \frac{\varepsilon}{3}$$

wird für alle n. Es bleibt also noch das mittlere Integral übrig. Wenn wir zeigen können, daß dieses Integral, in dem a und b bereits fest gewählt sind, für $n \rightarrow +\infty$ gegen Null konvergiert, also für hinreichend große n kleiner als $\frac{\varepsilon}{3}$ ist, so wird damit wegen (176) die

Behauptung, daß nämlich für hinreichend große n die Ungleichheit $|R_n(\lambda) - R(\lambda)| < \varepsilon$ gilt, bewiesen sein. Nun ist aber das mittlere Integral $= \int_a^b f_n(\mu)\, d\mu$, wenn $f_n(\mu) = \frac{|\sigma_n(\mu) - \sigma(\mu)|}{(p-\mu)^2 + q^2}$ gesetzt wird. Wegen (175) ist $|f_n(\mu)| \leqq \frac{2M}{q^2} =$ konst., und es gilt $f_n(\mu) \to 0$ bis auf eine Nullmenge gewiß, da wegen $\sigma_n [\to] \sigma$ die Grenzgleichung $\sigma_n(\mu) - \sigma(\mu) \to 0$ an allen Stetigkeitsstellen der nicht abnehmenden Funktion $\sigma(\mu)$, also mit eventueller Ausnahme einer abzählbaren Punktmenge sicher zutrifft. Setzt man daher $f(\mu) \equiv 0$, so gilt nach dem Satz, den wir nach Arzelà benannt haben (S. 90), $\int_a^b f_m\, d\mu \to \int_a^b f\, d\mu = 0$, w. z. b. w.

Es gilt übrigens $R_n(\lambda) \to R(\lambda)$ nicht nur für ein jedes nicht reelle λ, sondern auch gleichmäßig auf jedem Kreisbereich, der die reelle Achse nicht trifft. Mit Rücksicht auf den Satz von Vitali genügt es zu beweisen, daß die Funktionen $R_n(\lambda)$ auf jedem derartigen Kreisbereich gleichmäßig beschränkt sind. Dies wissen wir jedoch bereits von S. 92 her. Wir haben dort nämlich die Abschätzung des Resolventenintegrals mittels der Totalschwankung der Belegung ausgeführt, und die Totalschwankung von σ_n bleibt wegen (175) unterhalb einer von n unabhängigen Schranke. — Aus demselben Grunde gilt $R_n(\lambda) \to R(\lambda)$ gewiß, und zwar gleichmäßig, auch auf jedem die reelle Achse treffenden Kreisbereiche, der in einem etwas größeren Kreise liegt, welcher nur Konstanzstellen der σ_n enthält.

§ 48. Oberbereich und Unterbereich. Der Grommer-Hamburgersche Fundamentalsatz.

Wir lassen nun in (175) die Bedingung $\sigma_n [\to] \sigma$ fallen, und setzen etwa die Normierung $\sigma_n(+0) + \sigma_n(-0) = 0$ voraus, so daß

$$(177) \quad |\sigma_n(\mu)| \leqq \int_{-\infty}^{+\infty} |d\sigma_n(\mu)| \leqq M, \; d\sigma_n \geqq 0; \; \sigma_n(+0) + \sigma_n(-0) = 0$$

gilt. Wir setzen ferner voraus, daß es einen Kreis K gibt, auf welchem die $R_n(\lambda)$ gegen eine Grenzfunktion $F(\lambda)$ konvergieren, deren Darstellbarkeit als ein Resolventenintegral nicht vorausgesetzt wird. Dies wird sich vielmehr, mit Rücksicht auf den soeben bewiesenen Grommerschen Satz, von selbst ergeben. Wir beweisen nämlich, daß es ein σ mit $\sigma_n [\to] \sigma$ gibt. Es handelt sich also um eine Umkehrung des Grommerschen Satzes. — Der Beweis verläuft so. Wäre die

Umkehrung unrichtig, so würde es wegen (177) nach § 38 zwei Teilfolgen geben derart, daß $\sigma_{h_n}[\rightarrow]\sigma^*$, $\sigma_{j_n}[\rightarrow]\sigma^{**}$ gilt und σ^* nicht $[=]\,\sigma^{**}$ ist. — Bezeichnen R^* und R^{**} die zu σ^* bzw. σ^{**} gehörigen Resolventenintegrale, so gilt nach dem vorher bewiesenen Grommerschen Satz $R_{h_n}\rightarrow R^*$ und $R_{j_n}\rightarrow R^{**}$ z. B. in der oberen Halbebene. Nun strebt aber R_n auf K, also nach dem Vitalischen Satz überhaupt, gegen F. Erst recht gilt also $R_{h_n}\rightarrow F$, $R_{j_n}\rightarrow F$ und daher offenbar $R^* = F$, $R^{**} = F$, mithin $R^* = R^{**}$. Die Stieltjessche Umkehrformel (168) ergibt daraus

$$\sigma^*(\mu) - [\sigma^*(+0) + \sigma^*(-0)] = \sigma^{**}(\mu) - [\sigma^{**}(+0) + \sigma^{**}(-0)]$$

für alle Stetigkeitsstellen, und die beiden [ ] sind hier wegen der dritten Beziehung (177) gleich Null. Es bleibt also $\sigma^*[=]\,\sigma^{**}$, während sich vorhin auch das Gegenteil ergeben hat. Unsere Annahmen sind also widerspruchsvoll, w. z. b. w. — Es gilt demnach das folgende

Konvergenzkriterium von Grommer-Hamburger. Es sei für die Folge $\{\sigma_n(\mu)\}$ die Bedingung (177) erfüllt, und es sei $\{R_n(\lambda)\}$ die Folge der zugehörigen Resolventenintegrale. Der Grenzwert $\lim\limits_{n=+\infty} R_n(\lambda)$ ist dann etwa für alle nicht reelle λ dann und nur dann vorhanden, wenn es eine Grenzbelegung σ mit $\sigma_n[\rightarrow]\sigma$ gibt; und zwar gilt dann

$$(178)\qquad \lim_{n=+\infty}\int_{-\infty}^{+\infty}\frac{d\sigma_n(\mu)}{\lambda-\mu} = \int_{-\infty}^{+\infty}\frac{d\sigma(\mu)}{\lambda-\mu},$$

d.h. die beiden Grenzprozesse $n\rightarrow+\infty$, $\int\limits_{-\infty}^{+\infty}$ dürfen vertauscht werden. —

Die durch die Stieltjessche Umkehrformel ausgesprochene Eineindeutigkeit der Abbildung des „Oberbereiches“ der Belegungen auf den „Unterbereich“ der Resolventenintegrale besteht also auch für die sinngemäß festgelegten Grenzelemente. — Man beachte, daß, wenn auch $\sigma_n[\rightarrow]\sigma$ bzw. $R_n[\rightarrow]R$ nicht gilt, konvergente Teilfolgen wegen (177) nach S. 78 bzw. S. 83 gewiß vorhanden sind, und daß beidemal die zur Auswahlfolge gehörigen Grenzelemente nach S. 82 bzw. S. 83 von der Art der Auswahl dann und nur dann unabhängig sind, wenn das Auswahlverfahren überflüssig ist. — Die Gesamtheit aller nicht reellen λ können wir freilich in dem Grommerschen Kriterium mit Rücksicht auf den Vitalischen Satz z. B. auch durch ein offenes Intervall ersetzen, das keine Nichtkonstanzstelle der σ_n enthält oder nicht auf der reellen Achse liegt, usf.

§ 49. Differential und linearer Operator.

Durch den vorangehenden Fundamentalsatz wird die Korrespondenz des „Stieltjesschen Differentials“ $d\sigma(\mu)$ und des zugehörigen Resolventenintegrals $R(\lambda)$ restlos beschrieben. Sobald es sich aber darum handelt, *zwei* Resolventenintegrale miteinander zu vergleichen, so erweist sich dieser Apparat noch als zu grob. Um z. B. die Resolventenmatrizen unendlicher Hermiteschen Matrizen miteinander näher in Beziehung setzen und damit die Frage der unitären Äquivalenz beantworten zu können, muß man vielmehr nach Hellinger die Belegung $d\sigma(\mu)$ aus gewissen „getrennten“ Normalbestandteilen additiv aufbauen. Diese additive Trennung ist, wie wir sehen werden, der Natur der Sache nach nicht eindeutig, indem die einzelnen Summanden auf die verschiedensten Weisen „Normalbestandteile“ bilden können. Wir werden uns zunächst nur mit den einzelnen Normalbestandteilen beschäftigen und kommen auf die additive Trennung selbst erst bei den unendlichen Matrizen zurück. — Wir beginnen mit einer Zerlegungsaufgabe von „$d\sigma(\mu)$“, die ganz anders, nämlich nicht additiv ist, sondern wie folgt lautet:

$$d\sigma(\mu) = \frac{|d\tau(\mu)|^2}{d\varrho(\mu)} \tag{179}$$

(der Sinn dieser „Formel“ soll später präzisiert werden). Hierin ist $\sigma(\mu)$ eine gegebene, nicht abnehmende Funktion; die Nennerfunktion $\varrho(\mu)$ wird vorgeschrieben und ist ebenfalls nicht abnehmend; die Funktion $\tau(\mu)$ wird gesucht und soll sozusagen das infinitesimale Tempo des Anwachsens der Funktion $\sigma(\mu)$ in bezug auf die Basis $\varrho(\mu)$ beschreiben. Die Frage nach der Lösbarkeit von (179) bei gegebenem σ und ϱ bezeichnen wir als das zur Basis ϱ gehörige Hellingersche Zerlegungsproblem von σ. Diese Zerlegung ist von der additiven Trennung unabhängig und schließt diese in sich, in dem sie eben die Normalgestalt der einzelnen Summanden liefert. Die additive Trennung kann u. a. darum auf die verschiedensten Weisen herbeigeführt werden, weil für die einzelnen Summanden die Basis gewissermaßen willkürlich vorgeschrieben werden kann. Doch wird es sich in diesem Kapitel, wie gesagt, nur um (179) handeln. Wir werden dabei, da wir es nur mit Stieltjesschen Integrationen zu tun haben, annehmen können, daß $\sigma(\mu)$ von rechts stetig ist. Nun ist aber die Zerlegung des rein sprunghaften Bestandteiles $\hat{\sigma}$ trivial [vgl. (142)] und läßt sich übrigens, wie Hellinger und Toeplitz gelegentlich ihrer Untersuchung über Jacobische Matrizen bemerkt haben, mit derjenigen

von $\tilde{\sigma}$ unmittelbar koppeln. So können wir daher wegen (126) annehmen, daß $\sigma(\mu)$ durchweg stetig ist.

Leider muß ich, um die Rahmen dieser Schrift nicht zu zersprengen, auf ins einzelne gehende Beweise, die stark in der Richtung der reellen Funktionentheorie liegen, verzichten und mich auf einen gemeinverständlichen Bericht über die Probleme der Zerlegungstheorie beschränken.

Es ist unnötig zu betonen, daß der Ausdruck „Stieltjessches Differential" nicht in dem üblichen Sinne des Wortes Differential zu verstehen ist. Hat $\sigma(\mu)$ etwa eine stetige Ableitung, so ist $d\sigma(\mu)$ vermöge des Mittelwertsatzes der Differentialrechnung gewiß nicht sinnlos. Dann ist nämlich $\underset{a}{\Delta}\sigma(x) = \frac{d\sigma(\mu)}{d\mu}\underset{a}{\Delta}x + \varepsilon$, wobei der Fehler ε von einer höheren Ordnung als $\underset{a}{\Delta}x$ ist, indem $\varepsilon / \underset{a}{\Delta}x \to 0$, $a \to \mu$ sogar gleichmäßig gilt (vgl. hierzu S. 52). Doch braucht die Funktion σ, auch wenn sie stetig ist, nicht durchweg differentiierbar zu sein, und der Ausdruck „StieltjesschesDifferential" gewinnt erst etwa durch die Stieltjessche Umkehrformel einen klaren Sinn. Ist σ unstetig, so braucht „$d\sigma(\mu)$" nicht „unendlich klein" zu sein, es hat vielmehr, in Hinblick auf seinen Beitrag zu $\int f(\mu)\,d\sigma(\mu)$, den „Wert"

$$\sigma(\mu + 0) - \sigma(\mu - 0) = \underset{\mu}{\Delta}\sigma \neq 0.$$

Führt man die Verteilungsfunktion $\sigma(\mu) \equiv 0$, $\mu < \alpha$; $\sigma(\mu) \equiv 1$, $\mu \geqq \alpha$ ein, so erkennt man, daß der Stieltjessche Integralbegriff auch die Diracschen Deltafaltungen als ein mathematisches Äquivalent umfaßt und dabei begrifflich bequemer (und rechentechnisch wenigstens nicht *wesentlich* unbequemer) ist. Daß es so sein muß, geht grundsätzlich auch aus einem F. Rieszschen Satze der Funktionalrechnung hervor. Es bezeichne nämlich $\Lambda(f)$ einen distributiven Operator der für $0 \leqq \mu \leqq 1$ erklärten stetigen Funktionen $f(\mu)$. Es sei also zu jedem solchen f eine Zahl $\Lambda(f)$ derart zugeordnet, daß

$$\Lambda(c_1 f_1 + c_2 f_2) = c_1 \Lambda(f_1) + c_2 \Lambda(f_2)$$

gilt. Es möge ferner der Operator in dem Sinne stetig, also „linear" sein, daß, wenn das zu $0 \leqq \mu \leqq 1$ gehörige Maximum der Abweichung $|f_n(\mu) - f(\mu)|$ der beiden Funktionen f_n, f für $n \to \infty$ gegen Null konvergiert, auch $\Lambda(f_n) \to \Lambda(f)$ gilt. Der F. Rieszsche Satz, der übrigens aus den oben entwickelten Hellyschen Sätzen in einigen Zeilen folgt, besagt nun, daß es eine von f unabhängige Funktion von beschränkter Schwankung, σ, gibt derart, daß $\Lambda(f) = \int_0^1 f(\mu)\,d\sigma(\mu)$ wird. Und zwar

ist σ durch den Operator Λ in allen Stetigkeitsstellen im Innern von $[0, 1]$ und außerdem für $\mu = 0$ und $\mu = 1$ bis auf eine additive Konstante vollständig bestimmt. Man erkennt dies, indem man auf die C. Neumannsche Reihe

$$\sum_{n=0}^{\infty} \frac{\Lambda(\mu^n)}{\lambda^{n+1}} = \Lambda\left(\frac{1}{\lambda - \mu}\right) = R(\lambda)$$

die Stieltjessche Umkehrformel anwendet (Satz von Lerch). Bei unendlichen Intervallen ist freilich der Sachverhalt komplizierter. Vgl. Kap. VI.

§ 50. Das Hellingersche Zerlegungsproblem.

a) Die beiden Funktionenpaare.

Es seien $\tau(\mu)$, $\varrho(\mu)$ zwei für $-\infty < \mu < +\infty$ erklärte stetige Funktionen von beschränkter Schwankung, die für $\mu \to -\infty$ verschwinden; $\varrho(\mu)$ sei darüber hinaus monoton nicht abnehmend (also reellwertig). Es möge ferner eine (also nicht nur eine) stetige, monotone und beschränkte Hilfsfunktion $F(\mu)$ vorhanden sein derart, daß für alle Intervalle a: $\alpha \leqq \mu \leqq \beta$ die Ungleichheit

(180) $$|\underset{\mathrm{a}}{\Delta}\tau|^2 \leqq \underset{\mathrm{a}}{\Delta}\varrho \cdot \underset{\mathrm{a}}{\Delta}F \qquad (\underset{\mathrm{a}}{\Delta}\varrho \geqq 0, \underset{\mathrm{a}}{\Delta}F \geqq 0)$$

gilt. Wir sagen dann, das Funktionenpaar (τ, ϱ) bilde ein Hellingersches Funktionenpaar erster Art. Dieser Begriff ist offenbar asymmetrisch in τ, ϱ, da (ϱ, τ) kein Funktionenpaar erster Art zu sein braucht.

Sind $\varrho(\mu)$, $\sigma(\mu)$ zwei für $-\infty < \mu < +\infty$ erklärte, monoton nicht abnehmende, für $\mu \to -\infty$ verschwindende, stetige und beschränkte Funktionen, so sagen wir, $[\varrho, \sigma] = [\sigma, \varrho]$ bilde ein Hellingersches Funktionenpaar der zweiten Art. Diese Beziehung ist also symmetrisch.

b) Der Operator erster Art.

Es sei $\alpha = \alpha_0 < \alpha_1 < \alpha_2 < \dots < \alpha_{N-1} < \alpha_N = \beta$ eine Einteilung des Intervalles $\mathrm{a} = [\alpha, \beta]$, man setze $\mathrm{a}_\nu = [\alpha_{\nu-1}, \alpha_\nu]$ und man bilde die Summe

(181) $$\sum_{\nu=1}^{N} \frac{|\underset{\mathrm{a}_\nu}{\Delta}\tau|^2}{\underset{\mathrm{a}_\nu}{\Delta}\varrho} \qquad (\underset{\mathrm{a}_\nu}{\Delta}\varrho \geqq 0),$$

wobei (τ, ϱ) von der ersten Art ist, so daß $\tau(\mu)$ wegen (180) auf a gewiß konstant ist, wenn $\varrho(\mu)$ auf a konstant ist. Aus $\underset{\mathrm{a}_\nu}{\Delta}\varrho = 0$ folgt daher $\underset{\mathrm{a}_\nu}{\Delta}\tau = 0$. Ist in (181) der Nenner, also auch der Zähler $= 0$, so wollen wir den „Quotienten" gleich Null setzen, wodurch (181) in

allen Fällen eindeutig erklärt ist. Man überzeugt sich leicht mittels der Schwarzschen Ungleichung, daß wegen (180)

$$(182) \qquad 0 \leqq \frac{\left|\underset{a}{\Delta}\tau\right|^2}{\underset{a}{\Delta}\varrho} \leqq \sum_{\nu=1}^{N} \frac{\left|\underset{a_\nu}{\Delta}\tau\right|^2}{\underset{a_\nu}{\Delta}\varrho} \leqq \underset{a}{\Delta}F, \qquad \sum_{\nu=1}^{N}{}' a_\nu = a$$

gilt. Die Summe (181) verkleinert sich also bei Weiterteilung der a_ν gewiß nicht, bleibt jedoch unterhalb einer nur von a abhängigen Schranke, die wegen der Stetigkeit der Funktion F nach Null konvergiert, wenn a auf einen Punkt P zusammenschrumpft. Beachtet man noch, daß dies gleichmäßig gilt für alle P eines jeden endlichen Intervalles, auf welchem ja F gleichmäßig stetig ist, so erkennt man nach sinngemäßer Wiederholung einer aus den Elementen bekannten Schlußweise (vgl. S. 84), daß die Summe (181) in dem auf S. 47 präzisierten Sinne einem Grenzwert zustrebt, wenn die Teilungen von $a = [\alpha, \beta]$ unendlich fein werden. Man bezeichnet diesen von Hellinger eingeführten Grenzwert mit

$$(183) \qquad \int_{\alpha}^{\beta} \frac{|d\tau(\mu)|^2}{d\varrho(\mu)}.$$

Da die Summe (181) sich bei der Weiterteilung nicht verkleinert, so gilt wegen (180)

$$(184) \qquad 0 \leqq \frac{\left|\underset{a}{\Delta}\tau\right|^2}{\underset{a}{\Delta}\varrho} \leqq \int_{\alpha}^{\beta} \frac{|d\tau(\mu)|^2}{d\varrho(\mu)} \leqq \underset{a}{\Delta}F; \qquad a = [\alpha, \beta].$$

Da $F(\mu)$, also auch $\underset{a}{\Delta}F$ nach Voraussetzung unterhalb einer festen Schranke liegt, so bleibt (183) wegen (184) für $\alpha \to -\infty$ oder $\beta \to +\infty$ beschränkt und nimmt dabei nicht ab, strebt also endlichen Grenzwerten zu, wodurch (183) auch für $\alpha = -\infty$, $\beta = +\infty$ erklärt werden kann. Setzen wir

$$(185) \qquad \sigma(\mu) = \int_{-\infty}^{\mu} \frac{|d\tau(x)|^2}{d\varrho(x)}; \qquad \sigma(-\infty) = 0,$$

so ist der Ausdruck (183) der Stetigkeit zufolge offenbar $= \underset{a}{\Delta}\sigma$, und (180) geht über in

$$(186) \qquad \left|\underset{a}{\Delta}\tau\right|^2 \leqq \underset{a}{\Delta}\varrho \cdot \underset{a}{\Delta}\sigma, \qquad 0 \leqq \underset{a}{\Delta}\sigma \leqq \underset{a}{\Delta}F.$$

Demnach ist die mittels (185) erklärte Funktion σ derart, daß in (180) die an sich willkürliche Hilfsfunktion F gewiß $= \sigma$ gesetzt werden

kann, da σ wegen der zweiten Ungleichheit (186) stetig ist. Wegen derselben Ungleichheit gibt es keine Hilfsfunktion F, die von langsamerem Wachstum wäre als σ. Ist aber F vom selben Wachstum wie σ, so daß $\underset{\mathrm{a}}{\Delta}\,\sigma \equiv \underset{\mathrm{a}}{\Delta}\,F$ gilt, so muß, da σ und F nach Null konvergieren, wenn der untere Endpunkt von a ins negative Unendliche rückt, auch $\sigma(\mu) \equiv F(\mu)$ gelten. Die Funktion (185) kann also dahin charakterisiert werden, daß sie in (180) für (τ, σ) *die* „Hilfsfunktion vom langsamsten Wachstum" darstellt. Dieser Umstand erhebt die zwischen den Differentialen hingeschriebene „Gleichung"

(187) $$d\sigma(\mu) = \frac{|d\tau(\mu)|^2}{d\varrho(\mu)}$$

zu einer inhaltsvollen Aussage. Doch ist dabei folgendes zu beachten: Es ist uns (τ, ϱ) gegeben, wir leiten daraus σ vermöge (185) ab, und *dann* gilt (187).

c) Das Umkehrproblem.

Uns interessiert jedoch die umgekehrte Fragestellung, wobei σ und ϱ gegeben sind und τ gesucht wird. Diese Aufgabe bezeichnen wir als das Hellingersche Zerlegungsproblem von σ mit Bezug auf die Basis ϱ. Wir wollen also (187) nach τ auflösen. Genauer ist unter dem Hellingerschen Zerlegungsproblem folgendes zu verstehen. Es sind zwei nicht abnehmende, stetige, beschränkte, für $\mu \to -\infty$ verschwindende Funktionen σ, ϱ gegeben. Sie bilden also im Sinne von S. 108 ein Funktionenpaar der zweiten Art. Es wird eine Funktion $\tau(\mu)$ gesucht, welche die beiden folgenden Eigenschaften hat:

1) Das Funktionenpaar (τ, ϱ) ist von der ersten Art.

[Es gibt also stetige, beschränkte, für $\mu \to -\infty$ verschwindende, der Bedingung (180) Genüge leistende Funktionen $F(\mu)$. Folglich kann, wie wir wissen, die stetige, nicht abnehmende, für $\mu \to -\infty$ verschwindende Funktion

(188) $$\xi(\mu) = \int\limits_{-\infty}^{\mu} \frac{|d\tau(\mu)|^2}{d\varrho(\mu)}$$

gebildet werden; sie ist die Hilfsfunktion von dem langsamsten Wachstum.]

2) Darüber hinaus wird von $\tau(\mu)$ verlangt, daß

(189) $$\xi(\mu) \equiv \sigma(\mu)$$

wird, wobei $\sigma(\mu)$ die vorgegebene, zu zerlegende Funktion bedeutet,

während $\xi(\mu)$ durch (188) eindeutig erklärt ist. Diese Bedingung 2) verlangt also eine *Eigenschaft sui generis,* indem die Eigenschaft (189) den beiden, *bereits vorliegenden* Funktionen ξ, σ entweder zukommt oder nicht zukommt.

Dann und nur dann ist es inhaltsvoll zu sagen, daß die vorgelegte Funktion σ mit Bezug auf die vorgelegte Basis ϱ gemäß (187) zerlegt ist, oder daß (185) nach τ aufgelöst werden kann, wenn wir τ den *beiden* Bedingungen 1), 2) gemäß wählen können.

d) Die erste Hälfte des Umkehrproblems: der Operator zweiter Art.

Wir zeigen zunächst, daß 1), der erste Teil des Hellingerschen Umkehrproblems, immer gelöst werden kann. Es wird sich sodann darum handeln, ob es unter den Lösungen $\tau(\mu)$ der Teilaufgabe 1) immer auch eine solche gibt, welche der Bedingung 2) genügt.

Die Teilaufgabe 1) verlangt, zu einem vorgelegten Funktionenpaar $[\sigma, \varrho] = [\varrho, \sigma]$ der zweiten Art eine Funktion $\tau(\mu)$ derart anzugeben, daß (τ, ϱ) von der ersten Art ist. Diese Teilaufgabe beschäftigt sich mit σ überhaupt nicht. Doch wollen wir sogleich eine Lösung von 1) herleiten, in welche auch σ eingeht, um dadurch später auf 2) Bezug nehmen zu können.

Wir werden nach Hellinger bei jedem $[\sigma, \varrho]$ durch einen zu (181) analogen Prozeß einen zu (183) analogen Operator

(190) $$\int_\alpha^\beta \sqrt{d\sigma(\mu)\, d\varrho(\mu)} = \int_\alpha^\beta \sqrt{d\varrho(\mu)\, d\sigma(\mu)}$$

einführen, der auch für $\alpha = -\infty$, $\beta = +\infty$ einen Sinn hat. Setzen wir sodann

(191) $$\eta(\mu) = \int_{-\infty}^{\mu} \sqrt{d\sigma(\mu)\, d\varrho(\mu)},$$

so ist $\tau = \eta$ eine Lösung von 1).

Es sei $[\sigma, \varrho]$ von der zweiten Art, und es bezeichne $\mathfrak{a} = \sum_{\nu=1}^{N} \mathfrak{a}_\nu$ eine Teilung von $\mathfrak{a}$. Es gilt $\Delta\varrho \geqq 0$, $\Delta\sigma \geqq 0$. Unter $\sqrt{A}$ soll für $A \geqq 0$ immer die nicht negative Bestimmung der Wurzel verstanden werden. Aus der Schwarzschen Ungleichheit folgt für die zu (181) analogen Summen

(192) $$\sum_{\nu=1}^{N} \sqrt{\underset{\mathfrak{a}_\nu}{\Delta}\varrho}\, \sqrt{\underset{\mathfrak{a}_\nu}{\Delta}\sigma}$$

wegen $\sum_{\nu=1}^{n} \underset{a_\nu}{\Delta}\varrho = \underset{a}{\Delta}\varrho$ sofort

$$(193)\qquad 0 \leqq \sum_{\nu=1}^{N} \sqrt{\underset{a_\nu}{\Delta}\varrho \underset{a_\nu}{\Delta}\sigma} \leqq \sqrt{\underset{a}{\Delta}\varrho \underset{a}{\Delta}\sigma}; \qquad \sum_{\nu=1}^{n} a_\nu = a,$$

so daß die Summen (192), in genauem Gegensatz zu den Summen (181), bei der Weiterteilung nicht zunehmen können. Sie streben, aus den selben Gründen wie die Summen (181), bei unendlich fein werdender Teilung von $a = [\alpha, \beta]$ in dem auf S. 47 erklärten Sinne einer Grenze zu, die mit (190) bezeichnet wird. Dieses Integral nimmt nicht ab, wenn α abnimmt oder β wächst, bleibt jedoch, wie aus (193) mit $a = [-\infty, +\infty]$ wegen $\Delta\varrho \geqq 0$, $\Delta\sigma \geqq 0$ mit Rüchsicht auf die Beschränktheit von ϱ und σ folgt, unterhalb der festen Schranke

$$(194)\qquad \sqrt{[\varrho(+\infty) - \varrho(-\infty)]\,[\sigma(+\infty - \sigma(-\infty)]} < +\infty,$$

strebt also für $\alpha \to -\infty$, $\beta \to +\infty$ endlichen Grenzen zu. Da (192) bei der Weiterteilung nicht zunehmen kann, so ist wegen (193) auch der Ausdruck (190), für welchen wir gemäß (191) offenbar $\underset{a}{\Delta}\eta$ schreiben können, gewiß $\leqq$ als die obere Schranke in (193), d. h. es ist

$$(195)\qquad 0 \leqq \underset{a}{\Delta}\eta \leqq \sqrt{\underset{a}{\Delta}\varrho \underset{a}{\Delta}\sigma}.$$

Der Vergleich von (195) mit (180) ergibt, daß (η, ϱ) ein Funktionenpaar der ersten Art bildet (σ ist dabei die Hilfsfunktion), so daß $\tau = \eta$, wie behauptet, eine Lösung des Teiles 1) des Hellingerschen Zerlegungsproblems darstellt. Es sei η^* irgendeine andere, nicht abnehmende, für $\mu \to -\infty$ verschwindende Lösung, welche die Hilfsfunktion σ zuläßt, so daß für alle a

$$(196)\qquad 0 \leqq \underset{a}{\Delta}\eta^* \leqq \sqrt{\underset{a}{\Delta}\varrho \underset{a}{\Delta}\sigma},$$

also auch

$$(197)\qquad \sum_{\nu=1}^{N} \underset{a_\nu}{\Delta}\eta^* \leqq \sum_{\nu=1}^{N} \sqrt{\underset{a_\nu}{\Delta}\varrho \underset{a_\nu}{\Delta}\sigma}; \qquad \sum_{\nu=1}^{N} a_\nu = a$$

gilt. Die linke Seite ist offenbar $= \underset{a}{\Delta}\eta^*$, die rechte strebt bei unbegrenzt verfeinernden Teilung der Grenze (190) zu, die aber $= \underset{a}{\Delta}\eta$ ist, so daß $\underset{a}{\Delta}\eta^* \leqq \underset{a}{\Delta}\eta$ ausfällt. Gilt hierbei das Gleichheitszeichen stets, so folgt wie auf S. 110, daß $\eta^*(\mu) \equiv \eta(\mu)$ ist. Mithin läßt sich (191) dahin charakterisieren, daß sie unter den der Bedingung (196) genügenden, nicht abnehmenden, für $\mu \to -\infty$ verschwindenden Funktionen *die* „Funktion vom raschesten Wachstum" ist.

e) Die Normierbarkeit der zweiten Hälfte des Umkehrproblems.

Der Teil 1) des Zerlegungsproblems ist damit erledigt. Wir gehen zum Problem selbst über, indem wir 2) nicht mehr außer acht lassen, und müssen zunächst die folgende Bemerkung vorausschicken: Ist $\tau(\mu)$ eine Lösung des Zerlegungsproblems [d. h. von 1) *und* 2)], so ist auch $\tau_* = \int\limits_{-\infty}^{\mu} |\, d\tau(\mu)\,|$, d. i. die Totalschwankung, eine Lösung.

Denn zunächst ist $|\underset{a_\nu}{\Delta}\tau|$ der Betrag der Differenz der beiden Werte, die τ in den Endpunkten von a_ν annimmt, also gewiß $\leqq$ als die Totalschwankung auf a_ν, d. h. es gilt $|\underset{a_\nu}{\Delta}\tau| \leqq |\underset{a_\nu}{\Delta}\tau_*|$, also offenbar auch

$$(198) \qquad \int\limits_{-\infty}^{\mu} \frac{|\, d\tau(\mu)\,|^2}{d\varrho(\mu)} \leqq \int\limits_{-\infty}^{\mu} \frac{|\, d\tau_*(\mu)\,|^2}{d\varrho(\mu)},$$

sobald der Ausdruck rechterhand überhaupt einen Sinn hat, also gewiß dann, wenn es nicht „gleich $+\infty$“ ist, genauer, wenn (τ_*, ϱ) ein Funktionenpaar der ersten Art bildet, was jedoch, wie wir jetzt zeigen, immer zutrifft. Zunächst ist nämlich τ eine Lösung, also hat das Integral (185) einen Sinn, (τ, ϱ) ist ein Paar der ersten Art, und (180) gilt, wie wir wissen, auch noch für $F = \sigma$. Mithin gilt $\sum\limits_{\nu=1}^{N} |\underset{a_\nu}{\Delta}\tau| \leqq \sum\limits_{\nu=1}^{N} \sqrt{\underset{a_\nu}{\Delta}\varrho\, \underset{a_\nu}{\Delta}\sigma}$, also, der Schwarzschen Ungleichung zufolge, erst recht

$$\sum_{\nu=1}^{N} |\underset{a_\nu}{\Delta}\tau| \leqq \sqrt{\sum_{\nu=1}^{N} \underset{a_\nu}{\Delta}\varrho \sum_{\nu=1}^{N} \underset{a_\nu}{\Delta}\sigma}.$$

Die rechte Seite ist hierbei offenbar $= \sqrt{\underset{a}{\Delta}\varrho\, \underset{a}{\Delta}\sigma}$, die linke strebt bei unbegrenzt verfeinernden Teilung der Totalschwankung $\int\limits_{\alpha}^{\beta} |\, d\sigma(\mu)\,|$ zu; diese ist, da nach Definition $\tau_*(\mu) = \int\limits_{-\infty}^{\mu} |\, d\sigma(\mu)\,|$ gilt und da $\tau_*(\mu)$ stetig ist, wegen $[\alpha, \beta] = a$ nichts anderes als $\underset{a}{\Delta}\tau_*$; augenscheinlich ist also

$$(199) \qquad |\underset{a}{\Delta}\tau_*|^2 \leqq \underset{a}{\Delta}\varrho\, \underset{a}{\Delta}\sigma,$$

so daß (τ_*, ϱ), wie behauptet, von der ersten Art ist.

Aus (199) folgt

$$\sum_{\nu=1}^{N} \frac{|\underset{a_\nu}{\Delta}\tau_*|^2}{\underset{a_\nu}{\Delta}\varrho} \leqq \sum_{\nu=1}^{N} \underset{a_\nu}{\Delta}\sigma = \underset{a}{\Delta}\sigma,$$

also auch $\int\limits_{\alpha}^{\beta} \frac{|d\tau_*(\mu)|^2}{d\varrho(\mu)} \leqq \underset{a}{\Delta}\sigma$, mithin

$$(200)\quad \int\limits_{-\infty}^{\mu} \frac{|d\tau_*(x)|^2}{d\varrho(x)} \leqq \sigma(\mu) - \sigma(-\infty) = \sigma(\mu), \quad \text{wobei} \quad \sigma(\mu) = \int\limits_{-\infty}^{\mu} \frac{|d\tau_*(\mu)|^2}{d\varrho(\mu)}$$

ist, da τ nach Voraussetzung eine Lösung des Zerlegungsproblems gibt. Aus (200) folgt nun, daß in (198) nicht nur das Zeichen $\leqq$, sondern auch das Zeichen $\geqq$, also in Wirklichkeit nur $=$ gilt. Damit ist gezeigt, daß zugleich mit τ auch die Totalschwankung τ_* eine Lösung des Zerlegungsproblems darstellt.

f) Die Unlösbarkeit des Umkehrproblems.

Wenn also das Zerlegungsproblem überhaupt eine Lösung hat, so hat es auch eine nicht abnehmende Lösung.

Hellinger hat nun an Beispielen gezeigt, daß die Funktion (191) bei geeignetem $[\varrho, \sigma]$ sehr wohl identisch verschwinden kann, trotzdem dabei weder ϱ noch σ konstant ist[1]). Andererseits ist (191) diejenige, für $\mu \to -\infty$ verschwindende, nicht abnehmende Funktion, welche unter der Nebenbedingung (195) vom raschesten Wachstum ist. Diese Nebenbedingung ist aber für alle etwaigen Lösungen des Zerlegungsproblems erfüllt. Denn σ ist eine brauchbare Hilfsfunktion F in (180) [allerdings diejenige vom langsamsten Wachstum]. Es ist also sehr wohl möglich, daß das Zerlegungsproblem bei nicht identisch verschwindenden σ und ϱ höchstens die Funktion $\tau(\mu) \equiv 0$ als eine nicht abnehmende (für $\mu \to -\infty$ verschwindende) Lösung zulassen kann. Dann müssen aber auch diejenigen eventuellen Lösungen identisch verschwinden, die nicht nichtabnehmend sind. Denn ist τ eine Lösung, so ist τ_* eine nicht abnehmende Lösung, und aus $\tau_* \equiv 0$

[1]) Es liegt darin nichts erstaunliches, daß (191) identisch verschwinden kann, trotzdem die Glieder der Näherungssummen (192), deren Grenzwert (191) ist, durchweg > 0 sind. Man denke nur an das folgende diskontinuierliche Analogon, das, wenn es auch weniger tief liegt, in nuce die ganze Erscheinung enthält: Es ist sehr wohl möglich, daß

$$\lim_{n=\infty} \sum_{\nu=1}^{n} \left| \sqrt{A_\nu^{(n)} B_\nu^{(n)}} \right| = 0$$

ist, trotzdem jedes $A_\nu^{(n)}$ und $B_\nu^{(n)} > 0$ ausfällt. Um dies einzusehen, braucht man nur $A_\nu^{(n)} = B_\nu^{(n)} = \frac{1}{n^2}$, $\nu = 1, 2, \ldots, n$ zu setzen. — Die Sache gelingt also jetzt sogar dann, wenn $A_\nu^{(n)} = B_\nu^{(n)}$ ist; oben freilich nur dann, wenn man $\sigma \neq \varrho$ zuläßt.

folgt offenbar $\tau \equiv 0$. Es könnte also höchstens die Funktion $\tau \equiv 0$ eine Lösung sein. Sie ist aber keine Lösung, da sonst $\sigma(\mu)$ wegen (181), (185) identisch verschwinden würde, was nicht der Fall ist. Folglich ist das Zerlegungsproblem im allgemeinen unlösbar.

Des näheren ist der Sachverhalt der folgende. Hellinger hat gezeigt, daß, *wenn*

$$\sigma = \int_{-\infty}^{\mu} \frac{|d\tau|^2}{d\varrho}$$

nach τ aufgelöst werden kann, der Operator zweiter Art, $\tau = \int_{-\infty}^{\mu} \sqrt{d\varrho\, d\sigma}$, gewiß eine (nicht abnehmende) Lösung liefert. Doch ist die kursiv gesetzte Prämisse nicht immer erfüllt. Man erinnere sich hier an die bei 2) gemachte Bemerkung (S. 111), daß es sich um Eigenschaften sui generis handelt. — Der Algorithmus ist freilich verlockend, es liegt ja nahe, wie folgt schließen zu wollen: da die stetige monotone Funktion $\int_{-\infty}^{\mu} \sqrt{d\varrho\, d\sigma}$ nach den bei 1) bewiesenen immer gebildet werden kann, so folgt

$$(201) \qquad \int_{-\infty}^{\mu} \frac{\left| d \int_{-\infty}^{\mu} \sqrt{d\varrho\, d\sigma} \right|^2}{d\varrho(\mu)} = \int_{-\infty}^{\mu} \frac{\left|\sqrt{d\varrho\, d\sigma}\right|^2}{d\varrho} = \int_{-\infty}^{\mu} d\sigma = \sigma,$$

q. e. d. — Um dieses Vorgehen nicht nur einfach mit der allgemeinen Bemerkung abzutun, daß man mit Differentialen eigentlich nicht einmal dann rechnen sollte, wenn man es vorher fundiert hat, möge man folgendes bedenken: zwar ist (187) in dem auf S. 110 erklärten Sinne inhaltsvoll und ebenso ist die Beziehung

$$(202) \qquad d\eta(\mu) = \sqrt{d\varrho(\mu)\, d\sigma(\mu)}$$

mit Rücksicht auf die bei (196) bewiesene kennzeichnende Extremaleigenschaft von $\eta(\mu)$ sinnvoll; es handelt sich aber nicht um *dieselbe* Extremaleigenschaft, sondern einmal um ein minimales, das anderemal um ein maximales Wachstum (dem entspricht es, daß, wenn man von den Konstanzstrecken absieht, die Näherungssummen bei dem Prozeß der Weiterteilung für den ersten Operator nur wachsen, für den zweiten aber nur abnehmen — und u. U. auch bis zur Null abnehmen[1])). — Hellinger führt den Nachweis dafür, daß, *wenn*

$$(*) \qquad \sigma = \int_{-\infty}^{\mu} \frac{|d\tau|^2}{d\varrho}$$

---

[1]) Vgl. die vorige Fußnote.

eine Lösung τ hat, $\tau = \int\limits_{-\infty}^{\mu} \sqrt{d\sigma\, d\sigma}$ eine Lösung ist, dadurch, daß er zeigt, daß das Rechnen (201) in der *umgekehrten* Richtung ε-mäßig legalisiert werden kann. Genauer ausgedrückt: die Gleichung $\tau = \int\limits_{-\infty}^{\mu} \sqrt{d\varrho\, d\sigma}$ kann bei gegebenen τ und ϱ in bezug auf σ ohne Einschränkung aufgelöst werden, und zwar ist (*) gewiß eine Lösung. Doch gibt es, wie wir wissen, für dieses Umkehrproblem zweiter Art möglicherweise auch weitere nicht abnehmende Lösungen. Denn ist $\tau \equiv 0$, ohne daß $\varrho \equiv 0$ wäre, so bleibt σ in $\tau = \int\limits_{-\infty}^{\mu} \sqrt{d\varrho\, d\sigma}$ ganz und gar willkürlich. Die evtl. wesentlich mehrdeutige Lösbarkeit des zweiten Umkehrproblems bedingt so die eventuelle Unlösbarkeit des ersten.

g) Das Kriterium der Lösbarkeit.

Hellinger hat nun das Kriterium gefunden dafür, daß sein Zerlegungsproblem lösbar ist (es ist von jetzt an stets das eigentliche Umkehrproblem gemeint, bei welchem τ gesucht wird; das andere, immer lösbare Umkehrproblem ist für spektrale Zwecke nur ein Hilfsmittel und an sich wertlos). Es ist also ein Funktionenpaar zweiter Art, $[\varrho, \sigma] = [\sigma, \varrho]$, gegeben und es wird nach einer nicht abnehmenden Funktion τ gefragt, welche die Bedingungen 1), 2) von S. 110 befriedigt. Es bezeichne M_ϱ die Menge derjenigen Punkte auf der μ-Achse, in deren Umgebung die Basisfunktion $\varrho(\mu)$ nicht konstant ist. Es sei N irgendeine Teilmenge von M_ϱ, und $\mathsf{P}(\mathsf{N})$ diejenige Punktmenge auf der ϱ-Achse, auf welche N vermöge der Beziehung $\varrho = \varrho(\mu)$ abgebildet wird, endlich $\Sigma(\mathsf{N})$ das durch $\sigma = \sigma(\mu)$ entworfene Bild von N auf der σ-Achse. Damit es nun ein $\tau(\mu)$ von der verlangten Beschaffenheit gibt, ist, wie Hellinger bewiesen hat, notwendig und hinreichend, daß, sobald für irgendeine Teilmenge N von M_ϱ das ϱ-Bild $\mathsf{P}(\mathsf{N})$ eine Nullmenge ist, das σ-Bild $\Sigma(\mathsf{N})$ desselben N ebenfalls eine Nullmenge darstellt. (Die Bedingung fällt also, wie es mit Rücksicht auf (180) auch zu erwarten war, in ϱ, σ asymmetrisch aus.) Es ist dabei angenommen, daß $\mathsf{P}(\mathsf{M}_\varrho)$, das bei Matrizen einfach das Spektrum ist, ganz im Endlichen liegt, doch dürfte die Behandlung des allgemeinen Falles auf keinerlei grundsätzliche Schwierigkeit stoßen.

Betrachten wir den einfachsten Fall. Es sei $\sigma(\mu) \equiv 1$ für $\mu \geqq 1$, $\equiv 0$ für $\mu \leqq 0$, und stetig und ohne Konstanzstrecken wachsend für $0 \leqq \mu \leqq 1$, während $\varrho(\mu) \equiv 0$, $\equiv \mu$ oder $\equiv 1$ ist, je nachdem $\mu \leqq 0$,

$0 \leqq \mu \leqq 1$ oder $\mu \geqq 1$ gilt, so daß es sich um das Umkehrproblem

(203) $$\sigma(\mu) = \int_0^\mu \frac{|d\tau(\mu)|^2}{d\mu}$$

handelt, wobei $\tau(\mu)$ für $0 \leqq \mu \leqq 1$ gesucht wird. Hier ist M_ϱ das Einheitsintervall $[0,1]$, also darf N eine beliebige Teilmenge von $[0,1]$ sein. Ferner ist $\mathsf{P}(\mathsf{N})$ wegen $\varrho(\mu) = \mu$ mit N identisch. Das Hellingersche Kriterium verlangt also einfach, daß vermöge der Abbildung $\sigma = \sigma(\mu)$ jede Nullmenge des Intervalles $0 \leqq \mu \leqq 1$ in eine Nullmenge der σ-Gerade übergeht. Nun kann die reelle Funktionentheorie stetige, monotone Funktionen σ aufweisen, die nicht jede Nullmenge auf eine Nullmenge abbilden, trotzdem sie keine Konstanzstrecke haben. Man bezeichnet diese Funktionen als nicht „totalstetig (absolutstetig, absolument continue)". — Ist also σ nicht totalstetig, so ist das Umkehrproblem (203) unlösbar und umgekehrt.

Eine Wiedergabe des Beweises des Hellingerschen Kriteriums würde bereits im Spezialfalle (203) eine ausgiebige Verwendung der Borel-Lebesgueschen Maßtheorie erfordern. Wir müssen uns versagen, auf diesen Gegenstand näher einzugehen.

Der Zusammenhang des Hellingerschen Umkehrproblems mit dem gewöhnlichen möge dennoch angedeutet werden. Es handelt sich um die Frage, inwiefern die Differentiation als die zu der unbestimmten Integration inverse Operation betrachtet werden kann, bzw. umgekehrt.

h) Vergleich mit der Lebesgueschen Theorie.

Ist $\sigma(\mu)$ eine für $0 \leqq \mu \leqq 1$ erklärte stetig differentiierbare Funktion mit der Ableitung $\sigma'(\mu)$, so wird $\frac{d\sigma}{d\mu} = \sigma'(\mu)$ durch $\sigma = \int^\mu \sigma'(\mu)\, d\mu$ aufgelöst und umgekehrt, das Integral zunächst nur im Riemannschen Sinne verstanden. Ist die Ableitung $\sigma'(\mu)$ durchweg vorhanden und beschränkt, ohne stetig zu sein, so ist sie, wie Volterra vor vielen Jahren gezeigt hat, u. U. im Riemannschen Sinne nicht integrierbar, so daß das Umkehrproblem in diesem Rahmen nicht einmal in Angriff genommen werden kann. Hier haben nun Lebesgue, später Denjoy und auf eine mehr direkte Weise Perron eingegriffen. Uns interessiert nur der Fall, wenn σ durchweg stetig und monoton nichtabnehmend ist. Nach Lebesgue ist dann $\sigma'(\mu)$ mit evtl. Ausnahme einer μ-Nullmenge, die nicht abzählbar zu sein braucht, gewiß vorhanden; ist $\sigma'(\mu) = +\infty$ oder $= -\infty$, so wird μ nicht als eine Stelle der Differentiierbarkeit betrachtet, ist also ein Punkt der Null-

menge. Man setze $\chi(\mu) = \sigma'(\mu)$, wenn $\sigma'(\mu)$ in dem vorhin auseinandergesetzten Sinne vorhanden ist, und sonst, d. h. auf der Nullmenge, setze man $\chi(\mu)$ etwa $= 0$. Dies ist ganz belanglos, das Lebesguesche Integral ist ja Nullmengen gegenüber unempfindlich. Diese Funktion $\chi(\mu)$ hat im Sinne von Lebesgue stets ein unbestimmtes Integral $\theta(\mu) = \int^{\mu} \chi(\mu)\, d\mu$, also sogar dann, wenn sie nicht beschränkt ist. Dies entspricht gewissermaßen dem Umstand, daß der Teil 1) des Hellingerschen Umkehrproblems stets gelöst werden kann; dem Teile 2) entspricht ebenso die Frage, ob $\theta(\mu)$ mit $\sigma(\mu)$ mindestens in denjenigen Punkten identisch sein muß, wo $\sigma'(\mu)$ existiert, wo also $\chi(\mu)$ gewiß $\neq 0$, d. h. > 0 ist. Die Antwort ist wieder verneinend. Damit sie bejahend wird, muß $\sigma(\mu)$ eine Eigenschaft sui generis, diejenige der Totalstetigkeit haben. Das unbestimmte Integral $\theta(\mu)$ ist nämlich immer totalstetig. Ist umgekehrt σ totalstetig, so gilt $\sigma = \theta$ für alle μ, wo σ' vorhanden ist. Im allgemeinen ist jedoch θ nur der totalstetige Bestandteil in einer von Lebesgue und Carathéodory herrührenden additiven Zerlegung der stetigen monotonen Funktion σ. — Jede in der Gestalt eines unbestimmten Lebesgueschen Integrals darstellbare Funktion ist totalstetig. Also gibt es stetige monotone Funktionen $\sigma(\mu)$, bei welchen das Umkehrproblem $\sigma(\mu) = \int_0^{\mu} \chi(\mu)\, d\mu$ auch dann unlösbar ist, wenn man von den Punkten, wo $\sigma'(\mu)$ nicht existiert, absieht. Es ist übrigens eine formale Überlegenheit des Stieltjes-Hellingerschen Algorithmus gegenüber des Leibniz-Lebesgueschen, daß dieser eben nur „Differentiale“ kennt und den Begriff der Ableitung folgerichtig vermeidet (oder, wenn man will, preisgibt). Es wird dadurch erreicht, daß, wenn $\sigma(\mu)$ eine stetige, nicht abnehmende, für $\mu = 0$ verschwindende Funktion bezeichnet, die vier Integrale

$$\int_0^{\mu} d\sigma(\mu), \quad \int_0^{\mu} |\, d\sigma(\mu)\,|, \quad \int_0^{\mu} \frac{|\, d\sigma(\mu)\,|^2}{d\sigma(\mu)}, \quad \int_0^{\mu} \sqrt{d\sigma(\mu)\, d\sigma(\mu)},$$

wobei das erste im Stieltjesschen, die beiden letzten im Hellingerschen Sinne gemeint sind und das zweite die Totalschwankung bedeutet, wie aus den Definitionen hervorgeht, miteinander identisch sind. Wegen $\sigma(0) = 0$ und mit Rücksicht auf die Stetigkeit von σ sind sie also nach (146) gleich $\sigma(\mu)$. Hingegen ist die Beziehung $\int_0^{\mu} \sigma'(\mu)\, d\mu \equiv \sigma(\mu)$, nach den vorangehenden, bei nicht totalstetigem σ gewiß unrichtig.

Die Beziehungen des Operators (183) zu dem Lebesgueschen Integralbegriff sind bereits von Hellinger betrachtet worden. In seiner Dissertation, in der die vorstehende Theorie zur Lösung des Äquivalenzproblems unendlicher Hermiteschen Matrizen entwickelt wird, beweist Hellinger den Fischer-F. Rieszschen Satz, der alle im Lebesgueschen Sinne samt ihrem Quadrate integrierbaren Funktionen und nur diese umfaßt, unter Benutzung des Grenzprozesses erster Art. Die Bildungen (181) treten übrigens bereits in der einschlägigen Note von E. Fischer auf. Beziehungen zur Lebesgueschen Integrationstheorie waren Hellinger von Anfang an bekannt, er hat aber die Heranziehung dieser Theorie absichtlich vermieden, da ihm sein Apparat dem Spektrumproblem mehr angepaßt erschien (vgl. auch die Schlußbemerkung des vorigen Absatzes). In der Tat handelt es sich auch in der Matrizenmechanik um statistische Fragestellungen, deren sinngemäße Sprache von vornherein die Hellingersche sein muß.

Die restlose Aufdeckung des Verhältnisses des Hellingerschen Operators erster Art zu der Lebesgueschen Integration verdankt man H. Hahn, der (183) auf Lebesguesche Integrale geradezu zurückgeführt hat (nachdem entsprechendes für die Stieltjessche Integration Lebesgue selbst gelungen war). Die Reduktion gelingt unter Benutzung von gewissen inversen Funktionen, von welchen sich der Leser aus dem zu (203) gehörigen, vorher besprochenen Spezialfall des allgemeinen Hellingerschen Kriteriums, das von Hahn ebenfalls hergeleitet wurde, eine Vorstellung bilden kann. Vgl. auch S. 186 weiter unten. — Näheres Eingehen auf die Lebesguesche Theorie, die noch immer kein Gemeingut bildet, muß leider außerhalb des Rahmens dieser Schrift liegen.

i) Verallgemeinerungen.

Übrigens ordnen sich die Hellingerschen und Lebesgueschen Grenzprozesse einer sehr allgemeinen Theorie unter, nämlich derjenigen der absolut-additiven Mengenfunktionen, die, von Lebesgue unter einschränkenden Bedingungen begründet, von Radon weit entwickelt wurde. Wir müssen uns leider auch diesmal mit einem Hinweis begnügen.

Zu (183) müssen wir für spätere Zwecke noch zwei Bemerkungen machen, welche die Bildungen

$$(204) \qquad \int_{\alpha}^{\beta} \frac{d\varphi(\mu)\, d\psi(\mu)}{d\varrho(\mu)} \quad \text{bzw.} \quad \int_{\alpha}^{\beta} f(\mu) \frac{d\varphi(\mu)\, d\psi(\mu)}{d\varrho(\mu)}$$

betreffen.

Sind (τ_1, ϱ), (τ_2, ϱ) zwei Funktionenpaare der ersten Art mit der-

selben Basis ϱ (aber evtl. mit verschiedenen Hilfsfunktionen F), so streben die zu (181) analogen Summen

$$\sum_{\nu=1}^{N} \frac{\underset{\mathfrak{a}_\nu}{\Delta}\tau_1 \underset{\mathfrak{a}_\nu}{\Delta}\tau_2}{\underset{\mathfrak{a}_\nu}{\Delta}\varrho}$$

in dem auf S. 47 erklärten Sinne einem Grenzwert zu, der mit $\int\limits_\alpha^\beta \frac{d\tau_1\, d\tau_2}{d\varrho}$ bezeichnet wird. Offenbar ist

$$\int\limits_\alpha^\beta \frac{d\tau_1\, d\bar{\tau}_1}{d\varrho} = \int\limits_\alpha^\beta \frac{|\,d\tau_1\,|^2}{d\varrho},$$

und umgekehrt kann der neue Prozeß auf den alten zurückgeführt werden. Denn nach der Schwarzschen Ungleichung ist $(\tau_1 + \tau_2, \varrho)$ ein Funktionenpaar erster Art, und es gilt offenbar

$$(205)\quad \int\limits_\alpha^\beta \frac{d\bar{\tau}_1\, d\tau_2}{d\varrho} \pm \int\limits_\alpha^\beta \frac{d\tau_1\, d\bar{\tau}_2}{d\varrho} = \int\limits_\alpha^\beta \frac{|\,d(\tau_1 \pm \tau_2)\,|^2}{d\varrho} - \int\limits_\alpha^\beta \frac{|\,d\tau_1\,|^2}{d\varrho} - \int\limits_\alpha^\beta \frac{|\,d\tau_2\,|^2}{d\varrho},$$

da entsprechendes bereits für die Näherungssummen gilt. Es wird sich als bequem erweisen, (205) einfach als die Definitionsgleichung des ersten Operators (204) aufzufassen.

Unter

$$(206)\qquad \int\limits_\alpha^\beta f(\mu)\, \frac{|\,d\tau(\mu)\,|^2}{d\varrho(\mu)}$$

wollen wir einfach

$$(207)\qquad \int\limits_\alpha^\beta f(\mu)\, d\sigma(\mu), \qquad \sigma(\mu) = \int\limits_{-\infty}^{\mu} \frac{|\,d\tau(\nu)\,|^2}{d\varrho(\nu)}$$

verstehen. Es läßt sich übrigens leicht zeigen, daß (206) der Grenzwert der Näherungssummen

$$\sum_{\nu=1}^{N} f(\xi_\nu)\, \frac{|\underset{\mathfrak{a}_\nu}{\Delta}\tau\,|^2}{\underset{\mathfrak{a}_\nu}{\Delta}\varrho}$$

darstellt; dabei liegt ξ_ν auf $\mathfrak{a}_\nu$, f ist stetig, (τ, ϱ) ist von der ersten Art. — Daß der dadurch direkt erklärte Operator (206) auf Stieltjessche Integrale zurückgeführt werden kann, geht bereits aus dem F. Rieszschen Satz über lineare Funktionentransformation (S. 107) hervor; es handelt sich nur darum, daß die zugehörige Belegung, wie zu erwarten ist, eben durch das Hellingersche Integral (207) geliefert wird.

Drittes Kapitel.

Die beschränkten unendlichen Matrizen.

§ 51. Einleitung.

Unter einer unendlichen Matrix wollen wir ein quadratisches Schema verstehen, in welchem beide Zeiger von 1 bis $+\infty$ laufen. Die frühere Bezeichnung, $\mathfrak{A} = \| a_{ik} \|$, nehmen wir für unendliche Matrizen in Anspruch, während eine endliche Matrix von n^2 Elementen mit $\mathfrak{A}_{(n)} = \| a_{ik}^{(n)} \|$ bezeichnet werden soll (vgl. S. 1—2). Die Elemente von $\mathfrak{A}_{(n)}$ brauchen nur für $1 \leqq i, k \leqq n$ erklärt zu sein. Hingegen wollen wir unter $\mathfrak{A}_{[n]} = \| a_{ik} \|_{[n]}$ den n-ten Abschnitt der unendlichen Matrix $\mathfrak{A}$ verstehen; das ist diejenige Matrix mit n^2 Elementen, welche aus der unendlichen Matrix entsteht, wenn man alle Elemente fortläßt, deren mindestens ein Zeiger $> n$ ist. — Eine Matrix, die keinen in einer unteren Klammer stehenden Zeiger hat, ist fortan immer eine unendliche Matrix, und entsprechend ein Vektor $\mathfrak{x}$ ein Vektor mit unendlich vielen Komponenten, ferner E die Gesamtheit aller normierten Vektoren: $|\mathfrak{x}|^2 = \sum\limits_1^\infty |x_i|^2 = 1$, und F die Gesamtheit aller Paare $\mathfrak{x}, \mathfrak{y}$ von normierten Vektoren, während das, was früher E bzw. F war, jetzt mit $\mathsf{E}_{[n]}$, $\mathsf{F}_{[n]}$ bezeichnet wird. Wir schreiben ferner gelegentlich $\sum$ an Stelle von $\sum\limits_1^\infty$, also z. B.

$$\sum a_{ij} x_j = \sum_{j=1}^{\infty} a_{ij} x_j, \qquad \sum \left(\sum a_{ij} x_j\right) = \sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} a_{ij} x_j\right).$$

Leicht ersichtliche Bezeichnungen, wie z. B. $\overline{\mathfrak{A}}$, $\mathfrak{A}'$, $\mathfrak{A}^*$, Hermitescher Charakter ($\mathfrak{A}^* = \mathfrak{A}$), Gleichsein zweier Matrizen usw., übertragen wir unmittelbar auf unendliche Matrizen, ebenso die Addition endlich oder unendlich vieler Matrizen und ihre Multiplikation mit gewöhnlichen Zahlen. Entsprechend erklären wir das Produkt zweier Matrizen, $\mathfrak{A}\mathfrak{B} = \| \sum a_{ij} b_{jk} \|$, wobei jedoch zu beachten ist, daß die unendliche

Reihe, welche das (i, k)-te Element von $\mathfrak{A}\mathfrak{B}$ darstellt, divergieren, also $\mathfrak{A}\mathfrak{B}$ zunächst sinnlos sein kann.

Wir sagen, für $\mathfrak{A}$ sei die „Q-Bedingung" erfüllt und nennen $\mathfrak{A}$ eine Q-Matrix, wenn die Quadratsumme der Beträge der in der i-ten Zeile stehenden Elemente absolut konvergiert, $K_i = \sum_{k=1}^{\infty} |a_{ik}|^2 < +\infty$, und dies außerdem für $\mathfrak{A}'$, d. h. auch bei den Kolonnen gilt. Hingegen wird nicht vorausgesetzt, daß etwa K_i unterhalb einer von i unabhängigen Schranke bleibt, so daß also z. B. $\mathfrak{A} = \| k! \, \delta_{ik} \|$, wo $\mathfrak{E} = \| \delta_{ik} \|$ die unendliche Einheitsmatrix bezeichnet, eine Q-Matrix ist. Nimmt man in der zu einem endlichen n gehörigen Schwarzschen Ungleichung den Grenzübergang $n \to +\infty$ vor, so folgt

$$(\sum |\alpha_n \beta_n|)^2 \leqq \sum |\alpha_n|^2 \sum |\beta_n|^2 \leqq +\infty.$$

Sind daher beide Reihen $\sum \alpha_n^2$, $\sum \beta_n^2$ absolut konvergent, so ist es auch $\sum \alpha_n \beta_n$. Das Produkt zweier Q-Matrizen kann also gewiß und eindeutig gebildet werden, es besteht aus unbedingt konvergenten Reihen. In den Anwendungen kommen häufig Q-Matrizen vor. Dies hängt einerseits mit dem Fischer-Rieszschen Satz bzw. damit zusammen, daß es meist auf quadratische Mittelwerte ankommt, und andererseits damit, daß den spektralen Verhältnissen eben die Kugelmannigfaltigkeit (S. 46) angepaßt ist, welche mit Rücksicht auf die Umkehrung der Schwarzschen Ungleichheit das Bestehen der Q-Bedingung erzwingt.

In Verallgemeinerung der Hilbert-Toeplitzschen Untersuchungen über beschränkte Gleichungssysteme (s. unten) hat E. Schmidt außer den Q-Matrizen auch die noch allgemeineren Matrizen betrachtet, bei welchen nur $K_i < +\infty$, $i = 1, 2, \ldots$ vorausgesetzt wird. Ist $\mathfrak{A}$ Hermitesch, so ist freilich mit $K_i < +\infty$ auch die Q-Bedingung erfüllt. — In den letzten Jahren hat Carleman die Hilbertsche Theorie der beschränkten Hermiteschen Matrizen z. T. auf Hermitesche Q-Matrizen übertragen. Wir werden jedoch sehen, daß die Gültigkeit mancher Sätze der Beschränktheitstheorie für das Bestehen der Beschränktheitsbedingung geradezu charakteristisch ist. Eine abgerundete Theorie für nicht beschränkte Q-Matrizen scheint nicht möglich zu sein. Z. B. braucht das Produkt zweier Q-Matrizen nicht wieder eine Q-Matrix zu sein. Sind $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ Q-Matrizen, und nimmt man an, daß $\mathfrak{A}\mathfrak{B}$ und $\mathfrak{B}\mathfrak{C}$ ebenfalls Q-Matrizen sind, so kann dabei auch $(\mathfrak{A}\mathfrak{B})\mathfrak{C} \neq \mathfrak{A}(\mathfrak{B}\mathfrak{C})$ sein. Das (i, k)-te Element in diesen Matrizen ist

$$(208) \qquad A_{ik} = \sum_l \Big(\sum_j a_{ij} b_{jl} c_{lk}\Big) \quad \text{bzw.} \quad B_{ik} = \sum_j \Big(\sum_l a_{ij} b_{jl} c_{lk}\Big),$$

und es handelt sich darum, daß die Umordnung der beiden Reihen A_{ik}, B_{ik} ineinander u. U. nicht gestattet, d. h. $A_{ik} \neq B_{ik}$ ist. So stößt bereits die Transformation von Q-Matrizen auf Schwierigkeiten. Für alle Fälle wird sich jede Theorie der unendlichen Matrizen in dem von Hilbert begründeten methodischen Rahmen bewegen müssen: Die Möglichkeit einer Verallgemeinerung mancher Hilbertschen Sätze liegt in einigen Fällen sehr nahe, in anderen Fällen ist eine Verallgemeinerung überhaupt unmöglich. Auch kommt es vor allem darauf an, daß die Probleme in einer sinngemäßen Weise gestellt werden. Dies fehlte vor Hilbert, während die Fragestellungen seiner Beschränktheitstheorie auch bei anderen Matrizen wenigstens gestellt werden können; nur kann dabei die Antwort verneinend ausfallen, trotzdem sie bei beschränkten Matrizen bejahend ist. — Vor einiger Zeit habe ich die für die Quantenmechanik vor allem wichtige Klasse der „halbbeschränkten" Matrizen betrachtet; ich komme auf diese Klasse weiter unten etwas ausführlicher zurück. — Eine Spektraltheorie für die beschränkten nicht Hermiteschen Matrizen, die ich unlängst in Angriff genommen habe, will ich hingegen nur skizzieren und an der Hand der unitären Matrizen illustrieren.

§ 52. Definitionen.

Bezeichnet $\Phi(\mathfrak{A}_{[n]})$ die Kopplungsform, $\Psi(\mathfrak{A}_{[n]})$ die Bilinearform des n-ten Abschnittes von $\mathfrak{A}$, und entsprechend $\mathsf{M}(\mathfrak{A}_{[n]})$, $\mathsf{P}(\mathfrak{A}_{[n]})$ die obere Grenze von $|\Phi(\mathfrak{A}_{[n]})|$ bzw. $|\Psi(\mathfrak{A}_{[n]})|$ auf $\mathsf{E}_{[n]}$ bzw. auf $\mathsf{F}_{[n]}$, endlich $\mathsf{W}(\mathfrak{A}_{[n]})$ das durch die Abbildung

$$(209) \qquad \Phi = \Phi(\mathfrak{A}_{[n]}) \equiv \Phi(\mathfrak{A}_{[n]}; x_1, x_2, \ldots, x_n)$$

in die Φ-Ebene entworfene Bild von $\mathsf{E}_{[n]}$, so gilt

$$(210) \quad 0 \leqq \mathsf{M}(\mathfrak{A}_{[n]}) \leqq \mathsf{M}(\mathfrak{A}_{[n+1]}) \quad \text{und} \quad 0 \leqq \mathsf{P}(\mathfrak{A}_{[n]}) \leqq \mathsf{P}(\mathfrak{A}_{[n+1]}),$$

und jeder Punkt von $\mathsf{W}(\mathfrak{A}_{[n]})$ gehört auch $\mathsf{W}(\mathfrak{A}_{[n+1]})$ an. Dies folgt einfach daraus, daß, wenn $x_{n+1} \equiv 0$ vorausgesetzt wird, $\mathsf{E}_{[n+1]}$ in $\mathsf{E}_{[n]}$ und $\Phi(\mathfrak{A}_{[n+1]})$ in $\Phi(\mathfrak{A}_{[n]})$ übergeht, und entsprechendes gilt bei Ψ, F. Mit $\mathfrak{A}$ ist auch jedes $\mathfrak{A}_{[n]}$ von Hermiteschem Typus (und umgekehrt); ist also $\mathfrak{A}$ Hermitesch, so können wir auch die Zahlen $\mathsf{n}(\mathfrak{A}_{[n]})$, $\mathsf{m}(\mathfrak{A}_{[n]})$ betrachten, welche die untere bzw. obere Grenze von $\Phi(\mathfrak{A}_{[n]})$ — also nicht von $|\Phi(\mathfrak{A}_{[n]})|$ — darstellen. Aus dem vorher erwähnten Grunde ist

$$(211) \qquad \mathsf{n}(\mathfrak{A}_{[n+1]}) \leqq \mathsf{n}(\mathfrak{A}_{[n]}) \quad \text{und} \quad \mathsf{m}(\mathfrak{A}_{[n]}) \leqq \mathsf{m}(\mathfrak{A}_{[n+1]}).$$

Da die erwähnten Zahlenfolgen monoton sind, können wir die zu $n \to +\infty$

gehörigen Grenzwerte $\lim \mathsf{M}(\mathfrak{A}_{[n]})$, $\lim \mathsf{P}(\mathfrak{A}_{[n]})$, $\lim \mathsf{m}(\mathfrak{A}_{[n]})$, $\lim \mathsf{n}(\mathfrak{A}_{[n]})$ einführen. Sie sollen bzw. mit $\mathsf{M}(\mathfrak{A})$, $\mathsf{P}(\mathfrak{A})$, $\mathsf{m}(\mathfrak{A})$, $\mathsf{n}(\mathfrak{A})$ bezeichnet werden, wobei $\mathfrak{A}$ die unendliche Matrix bedeutet. Alle vier Grenzwerte können auch unendlich sein, die drei ersten $= +\infty$, der vierte $= -\infty$. Die beiden ersten sind $\geqq 0$ und nur für die Nullmatrix $= 0$. Die beiden letzten sind nur für Hermitesche $\mathfrak{A}$ erklärt und können auch $\leqq 0$ sein. Für alle Fälle ist

$$0 \leqq \mathsf{M}(\mathfrak{A}) \leqq \mathsf{P}(\mathfrak{A}) \leqq 2\,\mathsf{M}(\mathfrak{A}) \tag{212}$$

und, wenn $\mathfrak{A}$ Hermitesch ist,

$$\mathsf{M}(\mathfrak{A}) = \mathsf{P}(\mathfrak{A}) = \max(|\mathsf{n}(\mathfrak{A})|, |\mathsf{m}(\mathfrak{A})|), \quad \mathsf{n}(\mathfrak{A}) \leqq \mathsf{m}(\mathfrak{A}) \lesseqgtr 0. \tag{213}$$

Dies alles folgt daraus, daß die entsprechenden Beziehungen für die $\mathfrak{A}_{[n]}$, wie wir wissen, richtig sind. Ferner ist $\{\mathsf{W}(\mathfrak{A}_{[n]})\}$ eine Folge von ineinander geschachtelten konvexen Bereichen, wobei $|\Phi| \leqq \mathsf{P}(\mathfrak{A}_{[n]})$ der kleinste um den Nullpunkt geschlagene Kreis ist, der $\mathsf{W}(\mathfrak{A}_{[n]})$ enthält. Bezeichnet daher $\mathsf{W}(\mathfrak{A})$ den kleinsten abgeschlossenen Bereich, der alle $\mathsf{W}(\mathfrak{A}_{[n]})$ enthält, so gilt $|\Phi| \leqq \mathsf{M}(\mathfrak{A})$ für alle Punkte von $\mathsf{W}(\mathfrak{A})$, und es gibt Punkte, wo das Gleichheitszeichen gilt. Freilich kann $\mathsf{W}(\mathfrak{A})$ mit der Vollebene identisch sein oder sonst irgendwie ins Unendliche ragen. Ist dies nicht der Fall, so daß $\mathsf{P}(\mathfrak{A}) < +\infty$ oder, was dasselbe ist, $\mathsf{M}(\mathfrak{A}) < +\infty$ gilt, so heißt die Matrix $\mathfrak{A}$ *beschränkt*. Auf diese Matrizen bezieht sich die von Hilbert begründete, von Hellinger und Toeplitz weiter ausgebildete Theorie, von der vorher die Rede war. Es seien $\lambda_k^{(n)}$; $k = 1, 2, \ldots, n$ die Eigenwerte von $\mathfrak{A}_{[n]}$; $n = 1, 2, \ldots$. Ihre Gesamtheit nebst den Häufungsstellen möge Abschnittsspektrum von $\mathfrak{A}$ genannt werden. Offenbar (vgl. § 19) liegt diese Punktmenge in $\mathsf{W}(\mathfrak{A})$, also, wenn $\mathfrak{A}$ beschränkt ist, gewiß ganz im Endlichen. Die Umkehrung davon gilt im allgemeinen nicht; vgl. S. 37. Ist aber jedes $\mathfrak{A}_{[n]}$ normal, also z. B. Hermitesch, so folgt aus dem Toeplitzschen Satz auf S. 38 unmittelbar, daß $\mathfrak{A}$ dann und nur dann beschränkt ist, wenn das Abschnittsspektrum ganz im Endlichen liegt.

Eine (nicht notwendig beschränkte) Hermitesche Matrix $\mathfrak{A}$ nennen wir positiv definit, wenn $\mathsf{n}(\mathfrak{A}) > 0$ ist, und nichtnegativ definit, wenn nur $\mathsf{n}(\mathfrak{A}) \geqq 0$ bekannt ist. Daraus, daß jeder Abschnitt von $\mathfrak{A}_{[n]}$ im Sinne von § 15 positiv definit ist, folgt noch nicht, daß auch $\mathfrak{A}$ selbst es ist. Das Beispiel $\mathfrak{A} = \| \varepsilon_i \delta_{ik} \|$, wobei $0 < \varepsilon_i \to 0$, $i \to \infty$, also z. B. $\varepsilon_i = i^{-1}$ ist, belegt dies. Eine solche Matrix, bei welcher also $\mathsf{n}(\mathfrak{A}_{[n]}) > 0$, $n = 1, 2, \ldots$ und trotzdem $\mathsf{n}(\mathfrak{A}) = 0$ ausfällt, wollen wir als pseudodefinit bezeichnen.

§ 53. Der Hilbertsche Konvergenzsatz.

Wir beweisen zunächst den Hilbertschen Satz, wonach bei jedem beschränkten $\mathfrak{A}$ die Doppelreihe

$$(214)\qquad \Psi(\mathfrak{A}) = \Psi(\mathfrak{A}; \mathfrak{x}, \mathfrak{y}) = \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} a_{ik} x_i y_k$$

auf F im Sinne der Konvergenz der Doppelreihen konvergent ist, so daß es zu jedem $\varepsilon > 0$ ein Zahlenpaar (N, M) gibt derart, daß

$$(215)\qquad \left| \sum_{i=1}^{m'}\sum_{k=1}^{n'} a_{ik} x_i y_k - \sum_{i=1}^{m''}\sum_{k=1}^{n''} a_{ik} x_i y_k \right|$$

$$\text{für } m' > M, \quad m'' > M, \quad n' > N, \quad n'' > N$$

$< \varepsilon$ wird; hierbei sind $\mathfrak{x}$ und $\mathfrak{y}$ auf F fest gewählt, so daß keine gleichmäßige Konvergenz behauptet wird, was, wie das Beispiel der offenbar beschränkten Einheitsmatrix, $\mathfrak{A} = \mathfrak{E}$, zeigt, im allgemeinen auch nicht zutreffen würde. Es wird auch nicht behauptet, daß die Reihe (214) in jedem Punkte von F absolut konvergiert, was, wie das einfache Beispiel $a_{ik} = (i-k)^{-1}$, $i \neq k$; $a_{ii} = 0$ zeigt, ebenfalls nicht behauptet werden kann[1]). Einen Vektor $\mathfrak{x}$ oder die zugehörige Zahlenfolge $x_1, x_2, \ldots$ nennen wir quadratisch konvergent, wenn die Reihe $\sum x_i^2$ absolut konvergiert. Da die Doppelreihe $\Psi(\mathfrak{A})$ auf F konvergent ist, so ist sie aus Gründen der Homogenität gewiß konvergent, wenn $\mathfrak{x}$ und $\mathfrak{y}$ quadratisch konvergent sind. Auf F ist also die Bilinearform $\Psi(\mathfrak{A})$ erklärt. Setzen wir in (215) $m' = n'$, $m'' = n''$, so folgt $\Psi(\mathfrak{A}_{[n]}) \to \Psi(\mathfrak{A})$, $n \to +\infty$. Lassen wir zuerst N und dann M ins Unendliche wachsen oder umgekehrt, so folgt

$$(216)\qquad \Psi(\mathfrak{A}) = \sum_{i=1}^{\infty}\Big(\sum_{k=1}^{\infty} a_{ik} x_i y_k\Big) = \sum_{k=1}^{\infty}\Big(\sum_{i=1}^{\infty} a_{ik} x_i y_k\Big),$$

so daß man die Doppelreihe auf F auch zeilenweise, aber auch kolonnenweise summieren kann. Da die Doppelreihe, wie erwähnt, evtl. nicht in jedem Punkte von F absolut konvergent ist, so können freilich im allgemeinen die Terme nicht beliebig umgeordnet werden. Setzt man in Ψ den Vektor $\mathfrak{y}$ gleich $\mathfrak{x}$, so folgt, daß auf E die zu $\mathfrak{A}$ gehörige unendliche Kopplungsform $\Phi(\mathfrak{A})$ konvergent ist. Es gilt wieder $\Phi(\mathfrak{A}_{[n]}) \to \Phi(\mathfrak{A})$ usw. Das durch die Abbildung $\Phi = \Phi(\mathfrak{A})$ entworfene Bild von E ist daher, wenn noch die dem Bilde evtl. fehlenden Häufungspunkte hinzugefügt werden, mit der vorher durch W$(\mathfrak{A})$ bezeichneten

---

[1]) Vgl. § 111, S. 236.

Punktmenge identisch, die also als der Wertevorrat von $\mathfrak{A}$ (auf E) bezeichnet werden kann. Ist z. B. $\mathfrak{A} = \| k^{-1} \delta_{ik} \|$, so ist das Bild von E das von unten offene Intervall $0 < \Phi \leqq 1$, so daß dazu noch der Punkt $\Phi = 0$ hinzugefügt werden muß. Ist $\mathfrak{A}$ Hermitesch, so ist $\mathsf{m}(\mathfrak{A})$ die obere, $\mathsf{n}(\mathfrak{A})$ die untere Grenze von $\Phi(\mathfrak{A})$ auf E usw. — Es ist dabei überall folgendes zu beachten. Liegt $\mathfrak{x} = (x_1, x_2, \ldots, x_n, x_{n+1}, x_{n+2}, \ldots)$ auf E, so braucht $\mathfrak{x}_{[n]} = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$ noch nicht auf E zu liegen. Wegen

$$\Phi(\mathfrak{A}_{[n]}; \mathfrak{x}_{[n]}) \to \Phi(\mathfrak{A}; \mathfrak{x}), \tag{217}$$

und da $|\mathfrak{x}_{[n]}|^2 \to |\mathfrak{x}|^2 = 1$ gilt, bedeutet dies jedoch keine Schwierigkeit, weil andererseits $\mathsf{E}_{[n]}$ eine Teilmenge von E ist.

Um endlich die Konvergenz der Doppelreihe Ψ (auf F) nachzuweisen, nehmen wir etwa an, daß $m' \geqq m''$, $n' \geqq n''$ ist. Dann ist der Betrag der Differenz (215) gewiß

$$\leqq \left| \sum_{i=1}^{m''} \sum_{k=n''+1}^{n'} a_{ik} x_i y_k \right| + \left| \sum_{i=m''+1}^{m'} \sum_{k=1}^{n'} a_{ik} x_i y_k \right|. \tag{218}$$

Der erste dieser beiden Ausdrücke ist nach der Schwarzschen Ungleichung

$$\leqq \sqrt{\sum_{k=n''+1}^{n'} |y_k|^2} \sqrt{\sum_{k=n''+1}^{n'} \left| \sum_{i=1}^{m''} a_{ik} x_i \right|^2}. \tag{219}$$

Der erste Faktor ist für $n'' > N$, $n' > N$ offenbar $\leqq \sqrt{\sum_N^{\infty} |y_k|^2} < \frac{\varepsilon}{2\,\mathsf{P}(\mathfrak{A})}$, wenn N hinreichend groß ist. Der in (219) unter dem zweiten Wurzelzeichen stehende Ausdruck ist ein Teil aus der Kopplungsform einer Norm von $\mathfrak{A}_{[n']}$ oder von $\mathfrak{A}_{[m'']}$, je nachdem $n' \geqq$ oder $\leqq m''$ ist[1]). Es möge etwa der erste Fall vorliegen. Der unter dem Wurzelzeichen stehende Ausdruck ist dann $\leqq (\mathsf{P}(\mathfrak{A}_{[n']}))^2 \sum_{i=1}^{n'} |x_i|^2 \leqq (\mathsf{P}(\mathfrak{A}))^2$, da, wie wir wissen, $(\mathsf{P}(\mathfrak{A}_{[n]}))^2$ die obere Grenze der Normkopplungsformen auf $\mathsf{E}_{[n]}$ ist (S. 42). So ist der ganze Ausdruck (219) und damit der erste Ausdruck (218) gewiß $< \frac{\varepsilon}{2\,\mathsf{P}(\mathfrak{A})} \cdot \mathsf{P}(\mathfrak{A}) = \frac{\varepsilon}{2}$, und der zweite Ausdruck (218) wird durch passende Wahl von M ebenso $< \frac{\varepsilon}{2}$, also der ganze Ausdruck (218) offenbar $< \varepsilon$, w. z. b. w.

Wegen

$$\sum_{i=1}^{n} \left| \sum_{k=1}^{n} a_{ik} x_k \right|^2 \leqq (\mathsf{P}(\mathfrak{A}_{[n]}))^2, \quad \sum_{k=1}^{n} |x_k|^2 = 1 \text{ ist } \left| \sum_{k=1}^{n} a_{ik} x_k \right| \leqq \mathsf{P}(\mathfrak{A}),$$

[1]) Gewisse x sind dabei gleich Null zu setzen.

so daß $\sum_{k=1}^{n} |a_{ik}|^2$ mit Rücksicht auf die Umkehrung der Schwarzschen Ungleichung unterhalb einer von n und sogar auch von i unabhängigen Schranke bleibt. Da wegen $\mathsf{P}(\mathfrak{A}) = \mathsf{P}(\mathfrak{A}')$ zugleich mit $\| a_{ik} \|$ auch $\| a_{ki} \|$ beschränkt ist, so sind daher die beschränkten Matrizen gewiß Q-Matrizen.

Bevor wir weitergehen, müssen wir einige Bemerkungen über die unendlichen Linearformen vorausschicken. Ist in der Linearform $\mathsf{L}(\mathfrak{x}) = \sum \alpha_i x_i$ die Koeffizientenfolge $\{\alpha_n\}$ quadratisch konvergent, so ist diese Reihe auf E absolut konvergent und dem Betrage nach $\leqq \sqrt{\sum |\alpha_i|^2} = |\mathfrak{a}|$. Dies folgt aus der Schwarzschen Ungleichheit. Ist umgekehrt $\mathsf{L}(\mathfrak{x})$ in jedem Punkte von E konvergent, so hat man zugleich absolute Konvergenz, da einerseits E symmetrisch ist und andererseits jedes x_i nur in ein einziges Glied der Linearform eintritt. Mit Rücksicht auf die bei endlichen Linearformen gültige Umkehrung der Schwarzschen Ungleichung (S. 10) muß also die Reihe $\sum |\alpha_i|^2$ konvergieren. Im Punkte $x_i = \frac{\bar{\alpha}_i}{|\mathfrak{a}|}$ von E, wobei zunächst $|\mathfrak{a}| \neq 0$ sein möge (vgl. S. 10), ist die Linearform gleich $\sqrt{\sum |\alpha_i|^2}$. Es muß dann also $\{\alpha_i\}$ quadratisch konvergent sein und $|\mathfrak{a}|$ ist nicht nur eine obere Schranke und nicht nur die obere Grenze, sondern das angenommene Maximum von $|\mathsf{L}(\mathfrak{x})|$ auf E. Dies ist wieder die Umkehrung der Schwarzschen Ungleichheit. — Offenbar sind die Abschnitte $|\sum_{i=1}^{n} \alpha_i x_i|$ auf $\mathsf{E}_{[n]}$ dann und nur dann unterhalb einer von n unabhängigen Schranke gelegen, oder, wie wir wieder sagen können, die Linearform oder der Vektor ist dann und nur dann beschränkt, wenn $\sum |\alpha_i|^2$ konvergiert. Und zwar konvergiert dann $\sum \alpha_i x_i$ wegen $\sum_{N+1}^{\infty} |\alpha_i x_i| \leqq \sqrt{\sum_{N+1}^{\infty} |\alpha_i|^2} \to 0$ nicht nur absolut, sondern auch gleichmäßig auf E. Die Verhältnisse sind also wesentlich einfacher als bei den beschränkten Matrizen. Der Grund hierfür ist offenbar der folgende: in einer n-ären Linearform gibt es n Leerstellen (für die Variablen) und ebenso viele unabhängige Variable; bei einer n-ären Matrix gibt es n^2 Leerstellen, während die Bilinearform nicht n^2, sondern nur $2n$ Variable enthält, so daß eigentlich bereits in der Bilinearform eine Kopplung vorgenommen ist, die bei Vornahme des Grenzüberganges $n \to \infty$ die Komplikation der Verhältnisse bedingt[1]).

[1]) Vgl. die auf S. 46 erwähnte Arbeit von O. Toeplitz.

§ 54. Das Hellinger-Toeplitzsche Konvergenzkriterium.

Der Satz, daß aus der ausnahmslosen Konvergenz auf der Einheitsmannigfaltigkeit die Beschränktheit geschlossen werden kann, ist, wie Hellinger und Toeplitz bewiesen haben, auch für Matrizen richtig. M. a. W., die Beschränktheit von $\mathfrak{A}$ ist nicht nur hinreichend, sondern auch notwendig dafür, daß die Doppelreihe $\Psi(\mathfrak{A})$ in *jedem* Punkte von E konvergiert. Und zwar folgt die Beschränktheit von $\mathfrak{A}$ bereits daraus, daß Ψ auf F z. B. bei kolonnenweiser Summation konvergiert. Übrigens kann Ψ sogar durch Φ ersetzt werden, und z. B. daraus, daß $\Phi(\mathfrak{A}_{[n]}) \to \Phi(\mathfrak{A})$, wobei $\Phi(\mathfrak{A})$ eine eigentliche Zahl $(\neq \infty)$ bedeutet, in *jedem* Punkte von E gilt, folgt bereits die Beschränktheit von $\mathfrak{A}$. Doch ist der Übergang von Ψ zu Φ keine wesentliche Verschärfung. Vgl. nämlich (45).

Ist zunächst die kolonnenweise summierte Bilinearform

$$\Psi = \sum \left(\sum a_{ik} x_i\right) y_k \tag{220}$$

auf F konvergent, so muß die Reihe $\sum \left| \sum a_{ik} x_i \right|^2$ in jedem Punkte von E konvergieren. Dies folgt aus der Umkehrung der Schwarzschen Ungleichung mit $\alpha_k = \sum a_{ik} x_i$, wobei $\mathfrak{x}$ irgendeinen fest gewählten Punkt auf E bezeichnet. Es muß ferner, damit kolonnenweise die Summierung in Ψ auf E überhaupt ausgeführt werden kann, $\sum a_{ik} x_i$ auf E konvergieren, also $\sum\limits_i |a_{ik}|^2 < +\infty$ sein bei einem jeden i. Dies muß auch gelten, wenn $\sum \left| \sum a_{ik} x_i \right|^2$ in jedem Punkte von E konvergiert; und ist letztere Bedingung erfüllt, so ist die kolonnenweise summierte Reihe Ψ nach der Schwarzschen Ungleichung auf F gewiß konvergent. Die ausnahmslose kolonnenweise Konvergenz von Ψ auf F ist also gleichwertig mit der ausnahmslosen Konvergenz der reellen nicht negativen Reihe $H = \sum \left| \sum a_{ik} x_i \right|^2$ auf E, wobei es von vornherein feststeht, daß die inneren Summen auf E konvergent sein müssen. So läuft die erwähnte Hellinger-Toeplitzsche Umkehrung des Hilbertschen Satzes offenbar auf folgendes hinaus: Ist die *nichtnegative* Reihe H in jedem Punkte von E konvergent, so bleibt sie auf E unterhalb einer festen Schranke. Trivial ist auch der derart vereinfachte Satz nicht. Denn wegen der unendlichen Anzahl der Dimensionen kann etwa der Satz von Bolzano-Weierstraß zu einer Zurückführung ad absurdum nicht herangezogen werden. Den indirekten Beweis mußten vielmehr Hellinger und Toeplitz durch eine sinngemäße Anpassung des Cantorschen Diagonalprinzips führen. Des näheren muß auf ihre Arbeit in den Math. Annalen (Bd. 69, 1910) verwiesen werden, wo der

Beweis auch ohne einen expliziten Übergang zu H gelingt. — Wie F. Riesz bemerkt hat, handelt es sich bei der Angabe eines *festen* Punktes auf E, in dem die Reihe H divergieren muß, wenn sie auf E nicht beschränkt ist, um ein Analogon der Schlußweise, die z. B. den seit Lebesgue und Haar sehr geläufigen „diagonalen" Konstruktionen divergenter Fourierschen Reihen u. dgl. zugrunde liegt.

Die Reihe H ist offenbar das Quadrat des Betrages des Vektors $\mathfrak{A}'\mathfrak{x}$. Setzen wir also $\mathfrak{A}' = \mathfrak{B}$, so folgt, daß eine Matrix $\mathfrak{B}$ dann und nur dann beschränkt ist, wenn $|\mathfrak{B}\mathfrak{x}|^2$ für jeden normierten, d. h. der Bedingung $|\mathfrak{x}| = 1$ genügenden Vektor endlich ist. Nun kommt es aber auf den Wert von $|\mathfrak{x}|$ aus Gründen der Homogenität gar nicht an. Folglich ist eine Matrix $\mathfrak{B}$ dann und nur dann beschränkt, wenn sie jeden Vektor $\mathfrak{x}$ von endlichem Betrage in einen Vektor $\mathfrak{B}\mathfrak{x}$ von ebenfalls endlichem Betrage überführt.

§ 55. Der erste Hilbertsche Faltungssatz.

Wir wollen nun mit Hilbert aus seinem in § 53, S. 126 bewiesenen Konvergenzsatz einige grundlegende Folgerungen ziehen. Es sei $\mathfrak{A}$, $\mathfrak{B}$ ein Paar von beschränkten, also der Q-Bedingung gewiß genügenden Matrizen, so daß das Produkt $\mathfrak{A}\mathfrak{B} = \|\sum\limits_j a_{ij} b_{jk}\|$ eindeutig gebildet werden kann. Wir haben soeben bewiesen, daß die Reihe $\sum\limits_i |\sum\limits_k c_{ik} z_k|^2$ bei jeder beschränkten Matrix $\|c_{ik}\|$ für jede quadratisch konvergente Folge $\{z_k\}$ konvergiert. Da zugleich mit $\mathfrak{C}$ auch $\mathfrak{C}'$ beschränkt ist, so konvergiert daher die Reihe

$$(221) \qquad \Lambda = \sum_j \Big(\sum_i a_{ij} x_i\Big)\Big(\sum_k b_{jk} y_k\Big)$$

nach der Schwarzschen Ungleichung auf F gewiß. Sie ist

$$(222) \qquad = \Big(\sum_j \sum_i\Big) a_{ij} x_i \zeta_j, \qquad \zeta_j = \sum_k b_{jk} y_k,$$

unter $\Big(\sum\limits_j \sum\limits_i\Big)$ Summation im Sinne von (215) verstanden. Denn $\|a_{ij}\|$ ist beschränkt und die Folgen $\{x_i\}$, $\{\zeta_j\}$ sind quadratisch konvergent. Mit Rücksicht auf (216) ist daher

$$(223) \qquad \Lambda = \sum_i x_i \Big(\sum_j a_{ij} \zeta_j\Big) = \sum_i x_i \Big(\sum_j a_{ij} \Big(\sum_k b_{jk} y_k\Big)\Big).$$

Hierbei gilt

$$(224) \qquad \sum_j a_{ij} \Big(\sum_k b_{jk} y_k\Big) = \sum_k y_k \Big(\sum_j a_{ij} b_{jk}\Big).$$

Denn $\| b_{jk} \|$ ist beschränkt, i fest, also $\{ a_{ij} \}$ eine quadratisch konvergente Zahlenfolge und $\{ y_k \}$ eine andere, so daß (224) wieder nicht mehr aussagt, als daß die Bilinearform einer beschränkten Matrix bei quadratisch konvergenten Argumenten sowohl zeilenweise als auch spaltenweise summiert werden kann. Setzt man (224) in (223) ein, so bleibt wegen (222)

$$(225)\qquad \sum_j \left(\sum_i a_{ij} x_i\right)\left(\sum_k b_{jk} y_k\right) = \sum_i x_i \left(\sum_k y_k \left(\sum_j a_{ij} b_{jk}\right)\right).$$

Nun ist nach der Schwarzschen Ungleichung

$$(226)\qquad \left| \sum_j \left(\sum_i a_{ij} x_i\right) w_j \right| \leqq \sqrt{\sum_j \left| \sum_i a_{ij} x_i \right|^2},$$

für $\sum |w_j|^2 = 1$, und es gibt (bei festem $\{ x_i \}$) einen Punkt $\{ w_j \}$ mit $\sum |w_j|^2 = 1$ derart, daß bei $\leqq$ das Gleichheitszeichen gilt (nach der Umkehrung der Schwarzschen Ungleichung). Die obere Grenze von $\left| \sum_j \left(\sum_i a_{ij} x_i\right) w_j \right|$ für $\sum |x_i|^2 = \sum |w_j|^2 = 1$, m. a. W. die Zahl $\mathsf{P}(\mathfrak{A})$, ist daher derart, daß

$$(227)\qquad \sum_j \left| \sum_i a_{ij} x_i \right|^2 \leqq (\mathsf{P}(\mathfrak{A}))^2$$

in jedem Punkte von E gilt und daß diese Schranke nicht verbessert werden kann, d. h. genau die obere Grenze ist. Der in (225) linkerhand stehende Ausdruck ist nach der Schwarzschen Ungleichung absolut

$$(228)\qquad \leqq \sqrt{\sum_j \left| \sum_i a_{ij} x_i \right|^2} \sqrt{\sum_j \left| \sum_k b_{jk} y_k \right|^2} \leqq \sqrt{(\mathsf{P}(\mathfrak{A}))^2} \sqrt{(\mathsf{P}(\mathfrak{B}'))^2} = \mathsf{P}(\mathfrak{A}) \mathsf{P}(\mathfrak{B}).$$

Offenbar ist also das Produkt $\mathfrak{A}\mathfrak{B}$ beschränkt, aus (225) folgt daher unter Beachtung von (215)

$$(229)\qquad \sum_j \left(\sum_i a_{ij} x_i\right)\left(\sum_k b_{jk} y_k\right) = \left(\sum_i \sum_k\right) x_i y_k \sum_j a_{ij} b_{jk} \equiv \Psi(\mathfrak{A}\mathfrak{B}; \mathfrak{x}, \mathfrak{y}).$$

Dies ist der erste Hilbertsche Faltungssatz. Er besagt u. a. daß das Produkt zweier beschränkten Matrizen wieder beschränkt ist.

Es ist selbstverständlich, daß die Summe endlich vieler beschränkten Matrizen ebenfalls beschränkt ist. Aber auch die Summe von unendlich vielen beschränkten Matrizen $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \ldots$ ist gewiß beschränkt, wenn die Reihe

$$(230)\qquad \mathsf{P}(\mathfrak{A}) + \mathsf{P}(\mathfrak{B}) + \mathsf{P}(\mathfrak{C}) + \cdots \qquad (\geqq \mathsf{P}(\mathfrak{A} + \mathfrak{B} + \mathfrak{C} + \cdots))$$

konvergiert; und zwar ist dabei, da a_{ik} absolut offenbar $\leqq \mathsf{P}(\mathfrak{A})$ ist, die Reihe $a_{ik} + b_{ik} + c_{ik} + \cdots$, welche das (i, k)-te Element von $\mathfrak{A} + \mathfrak{B} + \mathfrak{C}$ darstellt, unbedingt konvergent. Aus (228) und (229) folgt

$$(231)\qquad \mathsf{P}(\mathfrak{A}\mathfrak{B}) \leqq \mathsf{P}(\mathfrak{A}) \mathsf{P}(\mathfrak{B}).$$

§ 56. Der zweite Hilbertsche Faltungssatz.

Eine Matrix $\mathfrak{A}$ nennt Toeplitz zeilenfinit, wenn in einer jeden (festen) Zeile nur endlich viele von Null verschiedene Elemente vorhanden sind. Ist außer $\mathfrak{A}$ auch $\mathfrak{A}'$ zeilenfinit, m. a. W., ist $\mathfrak{A}$ zugleich kolonnenfinit, so heißt $\mathfrak{A}$ finit. Die finiten Matrizen sind offenbar Q-Matrizen.

Die Behandlung des Resolventenproblems der Jacobischen Matrizen, die zu nicht „bestimmten" Stieltjesschen Momentenproblemen gehören, führt zu einem Matrizentripel $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ von der folgenden Beschaffenheit: 1) $\mathfrak{A}$ und $\mathfrak{C}$ sind beschränkt, $\mathfrak{B}$ ist finit, also sind alle drei Matrizen Q-Matrizen, so daß die Produkte $\mathfrak{A}\mathfrak{B}$, $\mathfrak{B}\mathfrak{C}$ eindeutig gebildet werden können; außerdem hat das Tripel die Eigenschaft, daß 2) $\mathfrak{A}\mathfrak{B}$ und $\mathfrak{B}\mathfrak{C}$ beschränkt sind. Da $\mathfrak{C}$ und $\mathfrak{A}$ ebenfalls beschränkt, also Q-Matrizen sind, können die Produkte $(\mathfrak{A}\mathfrak{B})\mathfrak{C}$, $\mathfrak{A}(\mathfrak{B}\mathfrak{C})$ eindeutig gebildet werden und sie sind nach dem ersten Faltungssatz sogar beschränkt; und das Tripel hat dennoch die Eigenschaft, daß 3) aus einem bereits auf S. 123 erwähnten Grunde $(\mathfrak{A}\mathfrak{B})\mathfrak{C} \neq \mathfrak{A}(\mathfrak{B}\mathfrak{C})$ ist. Allerdings ist dabei die mittlere (finite) Matrix $\mathfrak{B}$ nicht beschränkt. In der Algebra der beschränkten Matrizen gilt nämlich für die Matrizenmultiplikation das assoziative Gesetz:

Sind $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ beschränkte Matrizen, so sind nach dem ersten Faltungssatz die Matrizen $\mathfrak{A}\mathfrak{B}$, $\mathfrak{B}\mathfrak{C}$, $(\mathfrak{A}\mathfrak{B})\mathfrak{C}$, $\mathfrak{A}(\mathfrak{B}\mathfrak{C})$ gewiß beschränkt, und der zweite Faltungssatz, zu dem wir jetzt übergehen, besagt, daß $(\mathfrak{A}\mathfrak{B})\mathfrak{C} = \mathfrak{A}(\mathfrak{B}\mathfrak{C})$ gilt, ebenso wie bei den endlichen Matrizen.

Sind allgemeiner $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ solche Q-Matrizen, daß die Matrizen $\mathfrak{A}\mathfrak{B}$ und $\mathfrak{B}\mathfrak{C}$ und außerdem $\mathfrak{B}$ (also die *mittlere* Matrix) beschränkt sind, so gilt $(\mathfrak{A}\mathfrak{B})\mathfrak{C} = \mathfrak{A}(\mathfrak{B}\mathfrak{C})$. Die Behauptung besagt, daß bei einem jeden i und k

$$(232) \qquad \sum_l \Big(\sum_j a_{ij} b_{jl}\Big) c_{lk} = \sum_l a_{ij} \Big(\sum_l b_{jl} c_{lk}\Big)$$

gilt. Nun ist aber $\mathfrak{B} = \|b_{jl}\|$ beschränkt und die beiden Zahlenfolgen $x_j = a_{ij}$, $y_l = c_{lk}$ sind, da $\mathfrak{A}$ und $\mathfrak{C}$ Q-Matrizen sind, quadratisch konvergent (n. b., i und k sind fest), so daß (232) in (216) enthalten ist. Das Merkwürdige dabei ist es, daß die unbedingte Konvergenz der Reihen $\sum_j \sum_l a_{ij} b_{jl} c_{lk}$ nicht einmal dann behauptet werden kann, wenn alle drei Matrizen beschränkt sind. Denn setzt man $b_{jl} = (j-l)^{-1}$, $j \neq l$; $b_{jj} = 0$, so ist $\|b_{jl}\|$, wie bereits erwähnt (S. 125), beschränkt, und die Matrix $\|\,|b_{jl}|\,\|$ der Beträge nicht beschränkt, oder, was nach dem Hilbert-

schen Konvergenzsatz und seiner Hellinger-Toeplitzschen Umkehrung auf dasselbe hinausläuft, die Doppelreihe $\sum\sum b_{jl} x_j y_l$ ist auf F überall konvergent, doch gibt es Punkte auf F, wo sie nicht absolut konvergiert.

§ 57. Folgerungen. Das Majorantenkriterium der starken Konvergenz.

Da zugleich mit $\mathfrak{A}$ auch $\overline{\mathfrak{A}}$ und $\mathfrak{A}'$, also auch $\mathfrak{A}^*$ beschränkt ist, so sind wegen (231) die beiden Normen $\mathfrak{A}^*\mathfrak{A}$, $\mathfrak{A}\mathfrak{A}^*$ jeder beschränkten Matrix $\mathfrak{A}$ beschränkt, und wegen $(\mathfrak{A}^*\mathfrak{A})^* = \mathfrak{A}^*\mathfrak{A}^{**} = \mathfrak{A}^*\mathfrak{A}$, $(\mathfrak{A}\mathfrak{A}^*)^* = \mathfrak{A}\mathfrak{A}^*$ von Hermiteschem Typus. Setzt man in (229) $\mathfrak{B} = \mathfrak{A}^*$ und $y_k = \bar{x}_k$, so folgt

$$\sum_i \left| \sum_k a_{ik} x_k \right|^2 = \Phi(\mathfrak{A}\mathfrak{A}^*), \tag{233}$$

so daß $\mathsf{m}(\mathfrak{A}\mathfrak{A}^*) \geqq \mathsf{n}(\mathfrak{A}\mathfrak{A}^*) \geqq 0$, also wegen (213)

$$\mathsf{M}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{m}(\mathfrak{A}\mathfrak{A}^*), \qquad \mathsf{M}(\mathfrak{A}^*\mathfrak{A}) = \mathsf{m}(\mathfrak{A}^*\mathfrak{A}) \tag{234}$$

gilt. Aus (233) folgt mit Rücksicht auf das bei (227) Gesagte und wegen $\mathsf{P}(\mathfrak{A}^*) = \mathsf{P}(\mathfrak{A})$ offenbar

$$\mathsf{M}(\mathfrak{A}\mathfrak{A}^*) = (\mathsf{P}(\mathfrak{A}))^2 = \mathsf{M}(\mathfrak{A}^*\mathfrak{A}). \tag{235}$$

Es ergibt sich aus den bei jeder beschränkten Matrix gewiß richtigen Relationen (234), (235)

$$\mathsf{m}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{m}(\mathfrak{A}^*\mathfrak{A}) = (\mathsf{P}(\mathfrak{A}))^2. \tag{236}$$

Man könnte meinen, dies sei trivial. Denn es ist

$$\mathsf{m}(\mathfrak{A}_{[n]}\mathfrak{A}^*_{[n]}) = \mathsf{m}(\mathfrak{A}^*_{[n]}\mathfrak{A}_{[n]}) = (\mathsf{P}(\mathfrak{A}_{[n]}))^2 \tag{56$'$}$$

und nach Definition gilt $\mathsf{P}(\mathfrak{A}) = \lim \mathsf{P}(\mathfrak{A}_{[n]})$. So folgt jedoch nur

$$\lim_{n=\infty} \mathsf{m}(\mathfrak{A}_{[n]}\mathfrak{A}^*_{[n]}) = \lim_{n=\infty} \mathsf{m}(\mathfrak{A}^*_{[n]}\mathfrak{A}_{[n]}) = (\mathsf{P}(\mathfrak{A}))^2, \tag{237}$$

es müßte also die (wegen (236), (237) gewiß richtige) Grenzgleichung $\mathsf{m}(\mathfrak{A}_{[n]}\mathfrak{A}^*_{[n]}) \to \mathsf{m}(\mathfrak{A}\mathfrak{A}^*)$ direkt bewiesen werden. Nun ist die Richtigkeit dieser Grenzgleichung durchaus nicht selbstverständlich. Sonst könnte man wegen

$$\mathsf{n}(\mathfrak{A}_{[n]}\mathfrak{A}^*_{[n]}) = \mathsf{n}(\mathfrak{A}^*_{[n]}\mathfrak{A}_{[n]}) \tag{59$'$}$$

mit demselben Rechte schließen, daß außer (236) auch

$$\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{n}(\mathfrak{A}^*\mathfrak{A}) \tag{238}$$

gilt; (238) ist aber, wie wir sehen werden, nicht mehr bei jeder beschränkten Matrix $\mathfrak{A}$ richtig. Dem entspricht es, daß wir, als (59) für endliche Matrizen bewiesen wurde, auf S. 44 Determinantensätze herangezogen haben, die uns jetzt nicht mehr zur Verfügung stehen. Es handelt sich freilich nicht so sehr um den Apparat, vielmehr um Tatsachen, die davon unabhängig sind. Da es zwischen $\mathfrak{A}^* = \overline{\mathfrak{A}}'$ und $\mathfrak{A}'$, d. h. zwischen $\mathfrak{A}$ und $\overline{\mathfrak{A}}$ lösungstheoretisch keinen Unterschied gibt (man muß nur zu den konjugierten Größen übergehen), so läuft die Asymmetrie des Paares $(\mathfrak{A}, \mathfrak{A}^*)$ auf die Asymmetrie des Paares $(\mathfrak{A}, \mathfrak{A}')$ hinaus. Es wird sich also darum handeln, daß, im Gegensatz zu dem Hauptsatz über endliche Gleichungssysteme (S. 8), die zu $\mathfrak{A}$ und zu der transponierten Matrix $\mathfrak{A}'$ gehörigen linearen Gleichungssysteme bei beschränkten, sonst keinerlei Bedingung unterworfenen $\mathfrak{A}$ sich verschiedentlich verhalten können. Wir gehen darauf in den nächstfolgenden Paragraphen ausführlich ein. Zunächst wollen wir es klarmachen, *woher* es kommt, daß aus $\mathfrak{n}(\mathfrak{A}_{[n]}\mathfrak{A}_{[n]}^*) = \mathfrak{n}(\mathfrak{A}_{[n]}^*\mathfrak{A}_{[n]})$ durch einen Grenzübergang das Symmetriegesetz (238), das, wie erwähnt, bei beliebigen beschränkten Matrizen $\mathfrak{A}$ unrichtig ist, *nicht* geschlossen werden kann.

Einen Vektor $\mathfrak{a} = (\alpha_1, \alpha_2, \ldots)$ nennen wir beschränkt, wenn seine Norm $\sum |\alpha_i|^2$ endlich, d. h. wenn die zugehörige Linearform beschränkt ist. Eine Vektorfolge $\{\mathfrak{a}^{(m)}\}$ nennen wir gleichmäßig beschränkt, wenn die Normen unterhalb einer von m unabhängigen Schranke bleiben, $|\mathfrak{a}^{(m)}|^2 <$ konst. Hierfür ist hinreichend (jedoch nicht notwendig!), daß die Vektorfolge einen Majorantenvektor von endlicher Norm zuläßt, so daß es einen Vektor $\mathfrak{c} = (\gamma_1, \gamma_2, \ldots)$ mit $\gamma_i \geqq 0$ und $\sum \gamma_i^2 < +\infty$ gibt derart, daß $|\alpha_i^{(m)}| \leqq \gamma_i$ für alle i und m gilt. — Wir sagen, die Vektorfolge $\{\mathfrak{a}^{(m)}\}$ konvergiere *schwach* gegen einen Grenzvektor $\mathfrak{a}$ und schreiben hierfür $\mathfrak{a}^{(m)}(\rightarrow)\mathfrak{a}$, wenn die folgenden beiden Bedingungen erfüllt sind: I) Die Folge $\{\mathfrak{a}^{(m)}\}$ ist gleichmäßig beschränkt,

$$\sum_i |\alpha_i^{(m)}|^2 < K_{\mathfrak{a}} < +\infty, \qquad m = 1, 2, \ldots,$$

und II) es gilt bei jedem festen i die Grenzgleichung

$$\alpha_i^{(m)} \rightarrow \alpha_i, \quad m \rightarrow +\infty; \quad i = 1, 2, \ldots.$$

Offenbar ist dann der Grenzvektor $\mathfrak{a} = (\alpha_1, \alpha_2, \ldots)$ von einer Norm $\leqq K_{\mathfrak{a}} < +\infty$. Nun könnte man meinen, daß aus $\mathfrak{a}^{(m)}(\rightarrow)\mathfrak{a}$ und $\mathfrak{b}^{(m)}(\rightarrow)\mathfrak{b}$ sofort $\mathfrak{a}^{(m)}\mathfrak{b}^{(m)} \rightarrow \mathfrak{a}\mathfrak{b}$, d. h. $\sum\limits_i \alpha_i^{(m)}\beta_i^{(m)} \rightarrow \sum\limits_i \alpha_i\beta_i$, $m \rightarrow +\infty$

folgt (eine jede dieser Reihen ist der Schwarzschen Ungleichheit zufolge unbedingt konvergent und absolut $\leqq \sqrt{K_\mathfrak{a} K_\mathfrak{b}}$). Doch wird dies durch das folgende Beispiel widerlegt: $\alpha_i^{(m)} = \beta_i^{(m)} = \delta_{im}$, wobei $\| \delta_{ik} \| = \mathfrak{E}$ gesetzt ist. Hier gilt $\alpha_i^{(m)} \to 0$, $\beta_i^{(m)} \to 0$, $m \to \infty$ bei einem jeden i, so daß II) mit $\alpha_i = \beta_i = 0$, $\mathfrak{a} = \mathfrak{b} = \mathfrak{o}$ erfüllt ist. Aber auch I) ist erfüllt, und zwar mit $K_\mathfrak{a} = K_\mathfrak{b} = 1$. Trotzdem ist

$$\mathfrak{a}^{(m)} \mathfrak{b}^{(m)} = \sum_i \alpha_i^{(m)} \beta_i^{(m)} = 1, \quad \text{also} \quad \mathfrak{a}^{(m)} \mathfrak{b}^{(m)} \to 1 \neq 0 = \mathfrak{a}\mathfrak{b}.$$

Zur schwachen Konvergenz müssen wir also noch etwas hinzufügen. Wir beweisen das folgende, in gewissem Sinne nicht nur hinreichende, sondern auch notwendige

Majorantenkriterium der starken Konvergenz. Hat mindestens eine der beiden Folgen $\{\mathfrak{a}^{(m)}\}$, $\{\mathfrak{b}^{(m)}\}$ eine Majorante, so folgt $\mathfrak{a}^{(m)} \mathfrak{b}^{(m)} \to \mathfrak{a}\mathfrak{b}$ aus $\mathfrak{a}^{(m)} (\to) \mathfrak{a}$, $\mathfrak{b}^{(m)} (\to) \mathfrak{b}$ gewiß.

Korollar. Ist $\mathfrak{a}$ ein Vektor von endlicher Norm, so folgt $\mathfrak{a}\mathfrak{b}^{(m)} \to \mathfrak{a}\mathfrak{b}$ aus $\mathfrak{b}^{(m)} (\to) \mathfrak{b}$ gewiß [denn dann ist der Vektor $(|\alpha_1|, |\alpha_2|, \ldots)$ eine Majorante aller $\mathfrak{a}^{(m)} (\equiv \mathfrak{a})$].

Das Korollar besagt übrigens nur scheinbar weniger als das Majorantenkriterium, das in einigen Zeilen bewiesen werden kann. Es möge etwa $\{\mathfrak{a}^{(m)}\}$ eine Majorante $\mathfrak{c}$ haben, so daß $|\alpha_i^{(m)}| \leqq \gamma_i$, $\sum \gamma_i^2 < +\infty$, also für alle m

$$\sum_{i=N}^{+\infty} |\alpha_i^{(m)} \beta_i^{(m)}| \leqq \sqrt{\sum_{i=N}^{\infty} |\alpha_i^{(m)}|^2} \sqrt{\sum_{i=1}^{\infty} |\beta_i^{(m)}|^2} \leqq \sqrt{\sum_{i=N}^{\infty} \gamma_i^2} \sqrt{K_\mathfrak{b}} < \frac{\varepsilon}{3}$$

gilt, wenn N hinreichend groß ist. Da die beiden Reihen $\sum |\alpha_i|^2$, $\sum |\beta_i|^2$ konvergent, nämlich $\leqq K_\mathfrak{a}$ bzw. $K_\mathfrak{b}$ sind, so können wir N mit Rücksicht auf die Schwarzsche Ungleichung so groß wählen, daß

$$\sum_{i=N}^{+\infty} |\alpha_i \beta_i| < \frac{\varepsilon}{3}$$

wird. Nach erfolgter Wahl von N können wir ein m_0 derart festlegen, daß für $m > m_0$

$$\left| \sum_{i=1}^{N-1} \alpha_i^{(m)} \beta_i^{(m)} - \alpha_i \beta_i \right| < \frac{\varepsilon}{3}$$

ausfällt. Denn einerseits gilt wegen $\mathfrak{a}^{(m)} (\to) \mathfrak{a}$, $\mathfrak{b}^{(m)} (\to) \mathfrak{b}$ die Grenz-

gleichung $\alpha_i^{(m)} \beta_i^{(m)} \to \alpha_i \beta_i$ bei jedem i, und andererseits handelt es sich nunmehr nur um endlich viele (nämlich $N - 1$) Summanden.

Die Addition der drei letzten Abschätzungen ergibt

$$\left| \sum_{i=1}^{\infty} \alpha_i^{(m)} \beta_i^{(m)} - \sum_{i=1}^{\infty} \alpha_i \beta_i \right| < \varepsilon \quad \text{für} \quad m > m_0;$$

da $\varepsilon > 0$ beliebig klein sein kann, ist damit die Behauptung, $\mathfrak{a}^{(m)} \mathfrak{b}^{(m)} \to \mathfrak{a}\,\mathfrak{b}$, bewiesen.

§ 58. Das Reziprokenproblem. Die Iterationsreihe und der Hilbsche Kunstgriff.

Mit Rücksicht auf die beiden Faltungssätze kann man mit beschränkten Matrizen, sofern es sich um rationale ganze Operationen handelt, ebenso *rechnen* wie mit den endlichen Matrizen. Wir gehen nun zu den Untersuchungen von Toeplitz über das Reziprokenproblem über, welche zeigen, daß bei der Division wesentlich kompliziertere Verhältnisse vorliegen. Insbesondere sind vordere und hintere Reziproken voneinander wohl zu unterscheiden. In diesem Kapitel betrachten wir nur beschränkte Matrizen. Den weitertragenden Toeplitzschen Existenzbeweis (§ 111) werden wir dabei durch die Methode von Hilb ersetzen, die im Falle der Beschränktheit auf eine recht einfache Weise zum Ziele führt, indem sie in den C. Neumannschen Reihen, die der Beschränktheit zufolge ein Konvergenzgebiet haben werden, den Entwicklungsparameter einer projektiven Transformation unterwirft. Der Hilbsche Kunstgriff, der auch bei der Behandlung von weiteren komplizierten Funktionalgleichungen gute Dienste leistet, kommt übrigens implizite und in einem anderen Zusammenhang in dem Hermiteschen Fall bereits bei Hilbert vor.

Wir schicken die Definitionen voraus. $\mathfrak{B}$ heißt eine hintere Reziproke von $\mathfrak{A}$, wenn $\mathfrak{A}\mathfrak{B} = \mathfrak{E}$ gilt; $\mathfrak{A}$ heißt dabei vordere Reziproke von $\mathfrak{B}$. Alle Matrizen sind (bis Kap. VI) als beschränkt vorausgesetzt; dies soll später nicht immer wiederholt werden. Ist $\mathfrak{A}$ von Hermiteschem Typus, und gilt $\mathsf{n}(\mathfrak{A}) \geqq 0$ [die Definition von $\mathsf{n}(\mathfrak{A})$ s. auf S. 124], so heißt $\mathfrak{A}$ nichtnegativ definit; gilt genauer $\mathsf{n}(\mathfrak{A}) > 0$, so haben wir $\mathfrak{A}$ positiv definit genannt. Wegen (233) sind beide Normen jeder beschränkten Matrix nichtnegativ definite beschränkte Matrizen.

Wir beweisen zunächst den folgenden Hilfssatz von Hilb. Ist $\mathfrak{A}$ eine beschränkte Matrix mit $\mathsf{P}(\mathfrak{A}) < 1$, so konvergiert die nach den Iterationen fortschreitende E. Schmidtsche Reihe $\mathfrak{B} = \sum_{0}^{\infty} \mathfrak{A}^m$ und sie ist

sowohl vordere als auch hintere beschränkte Reziproke von $\mathfrak{E} - \mathfrak{A}$. Denn zunächst ist $\mathsf{P}(\mathfrak{A}^m) \leqq [\mathsf{P}(\mathfrak{A})]^m = \vartheta^m$, $0 < \vartheta < 1$, wegen (231). Mit Rücksicht auf das bei (230) Gesagte ist also $\mathfrak{B}$ eine beschränkte Matrix. Es ist

$$(\mathfrak{E} - \mathfrak{A})\,\mathfrak{B} = \sum_{m=0}^{\infty} \mathfrak{A}^m - \mathfrak{A} \sum_{m=0}^{\infty} \mathfrak{A}^m.$$

Da, wie man sich ohne Mühe überzeugt, $\mathfrak{A} \sum_{m=0}^{\infty} \mathfrak{A}^m = \sum_{m=0}^{\infty} \mathfrak{A}^{m+1}$ gilt, so folgt daraus $(\mathfrak{E} - \mathfrak{A})\,\mathfrak{B} = \mathfrak{A}^0 = \mathfrak{E}$. Ebenso verifiziert man auch $\mathfrak{B}(\mathfrak{A} - \mathfrak{E}) = \mathfrak{E}$, w. z. b. w.

Jetzt folgt der erwähnte projektive Kunstgriff, angewendet auf Hermitesche Matrizen. Der nicht Hermitesche Fall kann, wie wir im nächsten Paragraph sehen werden, auf den Hermiteschen zurückgeführt werden. — Es sei $\mathfrak{A}$ eine Hermitesche Matrix, für welche die Bedingung $\mathsf{M}(\mathfrak{A}) < 1$ nicht erfüllt zu sein braucht [vgl. (213)]. Es sei hingegen $\mathfrak{A}$ positiv definit. Auch dann gibt es eine beschränkte Matrix, die sowohl vordere als auch hintere Reziproke von $\mathfrak{A}$ ist. Zunächst ist nach Voraussetzung $\mathsf{n}(\mathfrak{A}) > 0$. Da $\Phi(\mathfrak{A})$ auf E durchweg $\geqq \mathsf{n}(\mathfrak{A}) > 0$ ist, so ist $|\Phi(\mathfrak{E} - \varkappa\mathfrak{A})| = |1 - \varkappa\,\Phi(\mathfrak{A})|$ auf E gewiß $\leqq 1 - \varkappa\,\mathsf{n}(\mathfrak{A}) < 1$, wenn die positive Zahl $\varkappa$ hinreichend klein ist. Wir legen sie irgendwie gemäß dieser Voraussetzung fest, so daß also $\mathsf{M}(\mathfrak{E} - \varkappa\mathfrak{A}) < 1$ gilt. Da zugleich mit $\mathfrak{A}$ auch $\mathfrak{E} - \varkappa\mathfrak{A}$ beschränkt ist, so gibt es wegen $\varkappa > 0$ nach (212) und nach dem im vorangehenden Absatz bewiesenen Hilfssatz eine beschränkte Matrix $\mathfrak{B}$, die sowohl vordere als auch hintere Reziproke von $\mathfrak{E} - (\mathfrak{E} - \varkappa\mathfrak{A}) = \varkappa\mathfrak{A}$ ist. Folglich ist $\varkappa\mathfrak{B}$ eine sowohl vordere als auch hintere beschränkte Reziproke von $\mathfrak{A}$, w. z. b. w.

§ 59. Das Reziprokenproblem. Die Toeplitzschen Kriterien und die Formalsätze.

Es sei $\mathfrak{A}$ eine beliebige, also nicht notwendig Hermitesche, beschränkte Matrix. Ihre hintere Norm, $\mathfrak{A}\mathfrak{A}^*$, ist entweder positiv definit oder nur nichtnegativ definit (ohne positiv definit zu sein). Es möge der erste Fall vorliegen, $\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) > 0$. Dann gibt es nach dem vorhergehenden eine beschränkte Matrix $\mathfrak{B}$ mit $(\mathfrak{A}\mathfrak{A}^*)\,\mathfrak{B} = \mathfrak{E}$. Mit Rücksicht auf die Faltungssätze besagt dies, daß $\mathfrak{A}$ mindestens eine hintere beschränkte Reziproke hat. Denn es ist ja $\mathfrak{A}(\mathfrak{A}^*\mathfrak{B}) = \mathfrak{E}$. — Ebenso beweist man, daß jede beschränkte Matrix $\mathfrak{A}$ mit einer positiv definiten vorderen Norm, $\mathsf{n}(\mathfrak{A}^*\mathfrak{A}) > 0$, mindestens eine vordere beschränkte Reziproke hat. — Nun können diese Sätze umgekehrt werden: $\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) > 0$

ist nicht nur hinreichend, sondern auch notwendig für die Existenz mindestens einer hinteren Reziproke; und ebenso folgt aus der Existenz mindestens einer vorderen Reziproke, daß $\mathsf{n}(\mathfrak{A}^*\mathfrak{A}) > 0$ gilt. Ist nämlich etwa die letztere Bedingung nicht erfüllt, so ist $\mathsf{n}(\mathfrak{A}^*\mathfrak{A}) = 0$. Zwecks Zurückführung ad absurdum nehmen wir mit Hellinger an, daß $\mathfrak{A}$ dennoch mindestens eine beschränkte vordere Reziproke $\mathfrak{C}$ hat. Wegen (229) und (233) gilt dann auf F

$$(239) \qquad 1 = \Psi(\mathfrak{E}) = \Psi(\mathfrak{C}\mathfrak{A}) = \sum_j \Big(\sum_i c_{ij} x_i\Big)\Big(\sum_k a_{jk} y_k\Big)$$
$$\leqq \sqrt{\Phi(\mathfrak{C}\mathfrak{C}^*;\mathfrak{x})}\,\sqrt{\Phi(\mathfrak{A}^*\mathfrak{A};\mathfrak{y})},$$

woraus wegen

$$0 \leqq \Phi(\mathfrak{C}\mathfrak{C}^*;\mathfrak{x}) \leqq \mathsf{M}(\mathfrak{C}\mathfrak{C}^*)$$

sofort

$$(239)' \qquad 1 \leqq \mathsf{M}(\mathfrak{C}\mathfrak{C}^*)\,\Phi(\mathfrak{A}^*\mathfrak{A};\mathfrak{y}),$$

folgt. Da $\mathsf{n}(\mathfrak{A}^*\mathfrak{A})$, mit Rücksicht auf die Definition $\mathsf{n}((\mathfrak{A}^*\mathfrak{A})_{[n]}) \to \mathsf{n}(\mathfrak{A}^*\mathfrak{A})$ und auf $\Phi((\mathfrak{A}^*\mathfrak{A})_{[n]}) \to \Phi(\mathfrak{A}^*\mathfrak{A})$, die genaue untere Schranke von $\Phi(\mathfrak{A}^*\mathfrak{A})$ auf E darstellt, so folgt daraus $1 \leqq \mathsf{M}(\mathfrak{C}\mathfrak{C}^*)\cdot\mathsf{n}(\mathfrak{A}^*\mathfrak{A})$, und daher wegen $\mathsf{n}(\mathfrak{A}^*\mathfrak{A}) = 0$ der Widerspruch $1 \leqq 0$. — Die Kriterien sind also, wie behauptet, notwendig und hinreichend.

Es gelten ferner nach Toeplitz bei beschränkten Matrizen die folgenden Formalsätze:

I) Hat $\mathfrak{A}$ mindestens eine vordere und mindestens eine hintere Reziproke, so sind sie miteinander identisch. Es gibt also dann offenbar nur eine vordere und nur eine hintere Reziproke, da ja z. B. alle vorderen Reziproken mit derselben hinteren Reziproke identisch sein müssen.

Der Beweis verläuft unter Benutzung der Faltungssätze wie folgt: Ist $\mathfrak{B}\mathfrak{A} = \mathfrak{E}$ und $\mathfrak{A}\mathfrak{C} = \mathfrak{E}$, so gilt $(\mathfrak{B}\mathfrak{A})\mathfrak{C} = \mathfrak{C}$ und $\mathfrak{B}(\mathfrak{A}\mathfrak{C}) = \mathfrak{B}$, mithin wegen $(\mathfrak{B}\mathfrak{A})\mathfrak{C} = \mathfrak{B}(\mathfrak{A}\mathfrak{C})$ auch $\mathfrak{C} = \mathfrak{B}$, w. z. b. w. Die Voraussetzung, daß alle drei Matrizen beschränkt sind, ist dabei sehr wesentlich: eine finite Matrix kann, wenn sie nicht beschränkt ist, unendlich viele verschiedene beschränkte hintere Reziproken haben, die alle auch vordere Reziproken sind (vgl. das auf S. 131 angedeutete Beispiel).

Den zweiten Formalsatz genügt es auszusprechen, da der Beweis, dank der Faltungssätze, von S. 6 wörtlich übernommen werden kann:

II) Hat die beschränkte Matrix $\mathfrak{A}$ eine und nur eine beschränkte hintere Reziproke, so ist sie zugleich eine, also wegen I) die einzige vordere Reziproke von $\mathfrak{A}$, und entsprechendes gilt für den Fall einer eindeutigen vorderen beschränkten Reziproke. —

Gibt es zwei verschiedene beschränkte hintere Reziproken, so gibt es deren sogar unendlich viele (und dabei wegen I) keine beschränkte vordere Reziproke) und entsprechendes gilt für vordere Reziproken. Denn es sei $\mathfrak{A}\mathfrak{B} = \mathfrak{E}$, $\mathfrak{A}\mathfrak{C} = \mathfrak{E}$ und $\mathfrak{B} \neq \mathfrak{C}$. Dann ist, wenn $\varkappa$ eine beliebige Zahl bedeutet, wegen

$$\mathfrak{A}(\mathfrak{B} + \varkappa\mathfrak{C} - \varkappa\mathfrak{B}) = \mathfrak{A}\mathfrak{B} + \varkappa\mathfrak{A}\mathfrak{C} - \varkappa\mathfrak{A}\mathfrak{B} = \mathfrak{A}\mathfrak{B} + \varkappa(\mathfrak{E} - \mathfrak{E}) = \mathfrak{E}$$

jede Matrix $\mathfrak{B} + \varkappa(\mathfrak{C} - \mathfrak{B}) \neq \mathfrak{B}$ eine beschränkte hintere Reziproke von $\mathfrak{A}$.

§ 60. Das Reziprokenproblem. Die Toeplitzsche Tabelle.

Die beiden Faltungssätze lassen also für eine beschränkte Matrix $\mathfrak{A}$ nur die vier folgenden Fälle zu:

(240)

1. Es gibt eine und nur eine hintere beschränkte Reziproke, die zugleich die einzige vordere beschränkte Reziproke ist.

2. Es gibt keine vordere und keine hintere beschränkte Reziproke.

3a. Es gibt keine beschränkte vordere Reziproke, es gibt aber unendlich viele hintere beschränkte Reziproken,

3b. oder es tritt das Umgekehrte ein.

Und zwar sind hierfür die notwendigen und hinreichenden Bedingungen nach den vorhergehenden bz. die folgenden:

(241)

1. $\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) > 0, \quad \mathsf{n}(\mathfrak{A}^*\mathfrak{A}) > 0.$

2. $\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) = 0, \quad \mathsf{n}(\mathfrak{A}^*\mathfrak{A}) = 0.$

3a. $\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) > 0, \quad \mathsf{n}(\mathfrak{A}^*\mathfrak{A}) = 0.$

3b. $\mathsf{n}(\mathfrak{A}^*\mathfrak{A}) > 0, \quad \mathsf{n}(\mathfrak{A}\mathfrak{A}^*) = 0.$

Ist $\{\beta_i\}$ eine quadratisch konvergente Zahlenfolge, so sind die beiden Matrizen

$$(242) \quad \mathfrak{A} = \left\| \begin{matrix} 0 & 1 & 0 & 0 \cdots \\ 0 & 0 & 1 & 0 \cdots \\ 0 & 0 & 0 & 1 \cdots \\ \vdots & & & \end{matrix} \right\|, \qquad \mathfrak{B} = \left\| \begin{matrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \cdots \\ 1 & 0 & 0 & 0 \cdots \\ 0 & 1 & 0 & 0 \cdots \\ & & \ddots & \end{matrix} \right\|$$

offenbar beschränkt, und, wie man leicht nachrechnet, derart, daß $\mathfrak{A}\mathfrak{B} = \mathfrak{E}$ gilt, wie auch dabei die β sein mögen. Für die Matrix $\mathfrak{A}$ liegt also der bei endlichen Matrizen noch nicht mögliche Fall **3a**

vor. Ist $\mathfrak{A}\mathfrak{B} = \mathfrak{E}$, so gilt $\mathfrak{A}'\mathfrak{B}' = \mathfrak{E}' = \mathfrak{E}$, so daß für eine Matrix $\mathfrak{A}$ der Fall **3a** dann und nur dann vorliegt, wenn für die transponierte Matrix der Fall **3b**, und umgekehrt. Ist $\mathfrak{A}$ normal, d. h. $\mathfrak{A}\mathfrak{A}^* = \mathfrak{A}^*\mathfrak{A}$, so ist der pathologische Fall **3** von vornherein ausgeschlossen (freilich ist dies nur eine hinreichende Bedingung). Insbesondere trifft dies bei den Hermiteschen Matrizen zu. Es ist übrigens zu beachten, daß daraus, daß $\mathfrak{A}$ normal ist, noch nicht folgt, daß auch $\mathfrak{A}_{[n]}\mathfrak{A}^*_{[n]} = \mathfrak{A}^*_{[n]}\mathfrak{A}_{[n]}$ gilt (die Umkehrung ist jedoch, mit Rücksicht auf das Majorantenkriterium auf S. 134, gewiß richtig). Hingegen ist $\mathfrak{A}$ dann und nur dann Hermitesch, wenn jedes $\mathfrak{A}_{[n]}$ Hermitesch ist.

Die Gleichberechtigung von $\mathfrak{A}$ und $\mathfrak{A}'$ bei endlichen Matrizen kann also ebenfalls keine gruppentheoretische Tatsache sein. In der Tat haben wir uns auf S. 8 auf eine sozusagen funktionentheoretische Tatsache gestützt, um dadurch, unter Benutzung der Formalsätze, den Fall **3** ausschließen zu können. Jetzt fehlt aber, wie wir sogleich sehen werden, die entsprechende Tatsache.

§ 61. Fortsetzung. Historische Bemerkungen.

Betrachten wir die zu einer beliebigen, nicht notwendig beschränkten Matrix $\mathfrak{A}$ gehörigen Gleichungen

$$(243) \qquad \sum_{k=1}^{\infty} a_{ik} x_k = c_i; \qquad i = 1, 2, \ldots.$$

Ist $\mathfrak{A}$ zeilenfinit, so sind die Reihen linkerhand bei jedem x_i-System konvergent. Freilich auch nur dann. Toeplitz hat gezeigt, daß auf diesen Fall diejenigen Sätze über endliche lineare Gleichungssysteme, die auch für Systeme gelten, bei welchen die Anzahl der Variablen mit derjenlgen der Gleichungen nicht übereinstimmt, dank des Gelingens eines „Endlichkeitsschlusses" übertragen werden können, was auch dabei die c_i sein mögen. Da die c_i keinerlei Konvergenz- oder Größenordnungsannahmen unterworfen sind, so läßt sich diesbezüglich auch betreffend der Lösungen gar nichts behaupten. — Wenn man nun von diesen Matrizen absieht und zu anderen abgerundeten Matrizenklassen übergeht, so ist eine *systematische* Theorie erst dann möglich, wenn man unter „Lösung" nur eine Lösung versteht, die ganz bestimmten, der Klasse angepaßten Konvergenzbedingungen genügt, und wenn man zugleich auch die zugelassenen c_i-Systeme passend einschränkt (trotzdem dabei für *gewisse* Matrizen der Klasse auch für *gewisse* andere c_i-Systeme Lösungen existieren, die der Konvergenzbedingung der Klasse genügen, oder aber bei *gewissen*, den Konvergenzbedingungen genügenden c_i-Systemen *gewisse* Matrizen der Klasse auch

Lösungen zulassen, die den Konvergenzbedingungen der Klasse nicht mehr genügen).

Diese Notwendigkeit hat gelegentlich bereits Poincaré erkannt, indem er in seinen beiden, sich unmittelbar nach der Hillschen Mondtheorie (1877—1878) datierenden, für die Theorie der unendlich vielen Variablen historisch so wichtig gewordenen Aufsätzen (1885—86) über diesen Gegenstand ausführt, daß ein System von unendlich vielen linearen Gleichungen gewissermaßen als ein ebensolches System von Ungleichungen aufzufassen ist (aus dieser Bemerkung von Poincaré ist methodisch z. B. die Borel-Denjoysche Theorie der quasianalytischen Funktionen entstanden; nur sind in der heutigen, von Carleman weit entwickelten Theorie die Spuren der Entstehung etwas verwischt, indem das Poincarésche Prinzip z. B. durch die [sich ihrerseits dem Poincaréschen Prinzip unterordnende] Bestimmtheitstheorie der asymptotischen Reihen ersetzt wird). — Daß dieses Prinzip zum Aufbau von Theorien nicht nur notwendig, sondern auch hinreichend ist, konnte aus verständlichen Gründen nicht sofort erkannt werden: es scheint, daß bei unendlichen Gleichungssystemen — sofern sie nicht von vornherein ganz harmloser Natur waren oder einer Behandlung mit einem *expliziten Apparat*[1]), etwa mit unendlichen Determinanten oder konvergenten Kettenbrüchen gestatteten — ähnliche Hemmungen vorlagen, wie gleichzeitig oder etwas früher bei den divergenten Reihen. So konnte das Prinzip, nach einer vorläufenden, durch die Enzyklopädie der Vergessenheit erst unlängst entrissenen Arbeit von Dixon (1902), nur bei Hilbert (1906) zur Durchführung gelangen. Hilbert legt dabei die an die Spektraltheorie angepaßte Bedingung der quadratischen Konvergenz, also die Kugelmannigfaltigkeit E zugrunde. Die Grundlagen eines allgemeinen, z. Z. nicht restlos entwickelten Rahmen sind in der auf S. 46 erwähnten Arbeit von Helly gelegt.

§ 62. Die lösungstheoretische Bedeutung der Reziproken.

In der Beschränktheitstheorie handelt es sich also zunächst[2]) um folgendes: In dem Gleichungssystem

$$(244) \qquad \mathfrak{A}\mathfrak{x} = \mathfrak{c}; \qquad \sum_{k=1}^{\infty} a_{ik} x_k = c_i, \quad i = 1, 2, \ldots$$

[1]) Wie unnatürlich dieser Standpunkt ist, geht daraus hervor, daß es außer der so darstellbaren Lösung auch weitere, an sich gleichberechtigte Lösungen geben kann. Ein „expliziter" Apparat liefert meistens diejenige Lösung, welche durch die Abschnitte festgelegt wird. (Vgl. Kap. VI.)

[2]) Vgl. hingegen § 97.

werden nur quadratisch konvergente $\mathfrak{c}$ zugelassen und als Lösung $\mathfrak{x}$ wird nur ein quadratisch konvergentes Zahlensystem $(x_1, x_2, \ldots)$ betrachtet, es wird also insbesondere der vom Nullvektor verschiedene Vektor $\mathfrak{x}$ nur dann als eine Eigenlösung der homogenen Gleichungen betrachtet, wenn die Norm von $\mathfrak{x}$ endlich ist, so daß es auch eine *normierte* Lösung gibt, z. B. $\frac{\mathfrak{x}}{|\mathfrak{x}|}$. Beachtet man nun, daß eine Matrix $\mathfrak{B} = \| b_{ik} \|$ dann und nur dann beschränkt ist, wenn die Reihe $\sum\limits_i | \sum\limits_k b_{ik} c_k |^2$ für alle quadratisch konvergenten Zahlenfolgen $\{c_k\}$ konvergiert (siehe S. 129), m. a. W., wenn sie jeden quadratisch konvergenten Vektor $\mathfrak{c}$ in einen ebensolchen, nämlich in $\mathfrak{B}\mathfrak{c}$ transformiert, so ersieht man unter Beachtung von (224) oder (229) durch Wiederholung der auf S. 5—6 durchgeführten Betrachtung, daß die beschränkten hinteren Reziproken von $\mathfrak{A}$ lösungstheoretisch gedeutet werden können: Die beschränkte Matrix $\mathfrak{A}$ hat dann und nur dann mindestens eine beschränkte hintere Reziproke $\mathfrak{B}$, wenn die Gleichungen $\mathfrak{A}\mathfrak{x} = \mathfrak{c}$ bei jedem quadratisch konvergenten $\mathfrak{c}$ mindestens eine quadratisch konvergente Lösung $\mathfrak{x}$ haben, und zwar ist $\mathfrak{x} = \mathfrak{B}\mathfrak{c}$ eine quadratisch konvergente Lösung; und es gibt bei *jedem* quadratisch konvergenten $\mathfrak{c}$ dann und nur dann genau eine quadratisch konvergente Lösung, wenn $\mathfrak{A}$ genau eine beschränkte hintere Reziproke hat. — Es sei betont, daß eine jede der Reihen $\sum\limits_k a_{ik} x_k$, $\sum\limits_k b_{ik} c_k$ absolut konvergiert, da $\mathfrak{A}$ und $\mathfrak{B}$, als beschränkte Matrizen, gewiß Q-Matrizen sind. — Die vorderen Reziproken sind, wie bereits erwähnt, die hinteren Reziproken der transponierten Matrizen und umgekehrt. Das Auftreten der pathologischen Fälle **3** erklärt sich daher auf die folgende Weise. Die zu der unter (242) stehenden, von Toeplitz angegebenen Matrix $\mathfrak{A}$ gehörigen inhomogenen Gleichungen sind $x_{i+1} = c_i$; $i = 1, 2, \ldots$. Diese Gleichungen sind bei jedem quadratisch konvergenten $\{c_i\}$ quadratisch konvergent lösbar, und zwar auf unendlichviele Weisen, da für x_1 von dem Gleichungssystem gar nichts verlangt wird. Die transponierte Matrix ist

(245)
$$\mathfrak{A}' = \mathfrak{A}^* = \left\| \begin{matrix} 0 & 0 & 0 & 0 \cdots \\ 1 & 0 & 0 & 0 \cdots \\ 0 & 1 & 0 & 0 \cdots \\ \vdots & & \ddots & \end{matrix} \right\|,$$

die zugehörigen Gleichungen sind $0 = c_1$, $x_i = c_{i+1}$, $i = 1, 2, \ldots$, also nicht ohne weiteres, sondern nur dann lösbar, wenn c_1 zufällig $= 0$ ist, so daß es keine hintere Reziproke, also für $\mathfrak{A}$ keine vordere

Reziproke geben kann. — Im Falle **2** sind drei Unterfälle denkbar: die Anzahl der linear unabhängigen normierten Lösungen des homogenen Gleichungssystems $\mathfrak{A}\mathfrak{x} = \mathfrak{o}$ ist entweder **2a.** gleich Null, oder **2b.** von Null verschieden, jedoch endlich, oder **2c.** unendlich. — Bei den endlichen Matrizen ist Fall **2c** ausgeschlossen, und mit Rücksicht auf den Alternativsatz auch **2a**. Bei unendlichen beschränkten und zwar auch bei Hermiteschen Matrizen kann, wie wir sehen werden, jeder der drei Unterfälle vorkommen. Der Alternativsatz ist also in diesem Rahmen unrichtig. Liegt Fall **2b** vor, so haben, wie nicht näher ausgeführt werden soll, die inhomogenen Gleichungen, ebenso wie bei endlichen Matrizen, bei gewissen, jedoch nicht bei allen quadratisch konvergenten $\mathfrak{c}$ eine, freilich nicht *nur* eine quadratisch konvergente Lösung (denn zur Lösung des inhomogenen Systems kann eine Eigenlösung addiert werden). — Im Falle **1** gibt es keine normierte Eigenlösung, da die Lösung bei jedem $\mathfrak{c}$ eindeutig bestimmt ist und der Nullvektor $\mathfrak{o}$ eine zu $\mathfrak{c} = \mathfrak{o}$ gehörige, nicht normierbare Lösung darstellt. — Im Falle **3a** gibt es bei jedem $\mathfrak{c} \neq \mathfrak{o}$ unendlich viele verschiedene Lösungen, also haben auch die homogenen Gleichungen normierte Lösungen. Denn sind $\mathfrak{x}_1, \mathfrak{x}_2$ zwei, zum selben $\mathfrak{c}$ gehörige verschiedene Lösungen, so ist $\mathfrak{x}_1 - \mathfrak{x}_2$, also auch $\frac{\mathfrak{x}_1 - \mathfrak{x}_2}{|\mathfrak{x}_1 - \mathfrak{x}_2|}$ eine Lösung der homogenen Gleichungen. Der Alternativsatz ist also jetzt ebenfalls ungültig, jedoch in einem umgekehrten Sinne wie bei **2a**. — Im Falle **3b** kann, wie das Beispiel (245) zeigt, der Alternativsatz in demselben Sinne wie bei **2a** ungültig sein.

§ 63. Die Resolvente. Existenzsatz.

Wir gehen nun zur Resolvente der beschränkten Matrizen über. Bezeichnet λ eine Zahl, so ist zugleich mit $\mathfrak{A}$ auch $\lambda\mathfrak{E} - \mathfrak{A}$ beschränkt. Die Punkte der λ-Ebene, in welchen für $\lambda\mathfrak{E} - \mathfrak{A}$ nicht der Fall **1** vorliegt, bezeichnen wir als die singulären Stellen der Matrix $\mathfrak{A}$, die übrigen als die regulären Stellen, so daß λ für $\mathfrak{A}$ dann und nur dann regulär ist, wenn $\lambda\mathfrak{E} - \mathfrak{A}$ genau eine beschränkte Reziproke

$$(\lambda\mathfrak{E} - \mathfrak{A})^{-1} = \| R_{ik}(\lambda) \|$$

hat, die, als Funktion von λ aufgefaßt, wieder als Resolvente von $\mathfrak{A}$ bezeichnet wird. Da $(\lambda\mathfrak{E} - \mathfrak{A})^{-1}$ nach Voraussetzung sowohl hintere als auch vordere Reziproke von $(\lambda\mathfrak{E} - \mathfrak{A})$ ist, so ist sie mit $(\lambda\mathfrak{E} - \mathfrak{A})$ vertauschbar und es liegt auch für $(\lambda\mathfrak{E} - \mathfrak{A})^{-1}$ der Fall **1** vor. Sind andererseits $\mathfrak{B}$, $\mathfrak{C}$ zwei beschränkte Matrizen, für welche der Fall **1** vorliegt, so liegt auch für das Produkt $\mathfrak{B}\mathfrak{C}$ der Fall **1** vor, da unter Beachtung der Faltungssätze die beschränkte Matrix $\mathfrak{C}^{-1}\mathfrak{B}^{-1}$ sowohl

vordere als auch hintere Reziproke von $\mathfrak{B}\,\mathfrak{C}$ ist. Sind dabei $\mathfrak{B}$ und $\mathfrak{C}$ vertauschbar, so sind es auch $\mathfrak{B}^{-1}$ und $\mathfrak{C}^{-1}$. Unter wiederholter Beachtung der Faltungssätze folgt daraus nach wortgetreuer Wiederholung der bei endlichen Matrizen benutzten Schlußweise (S. 16) die Hilbertsche Funktionalgleichung der Resolvente $\mathfrak{R}_\lambda = (\lambda\,\mathfrak{E} - \mathfrak{A})^{-1}$: sind λ', λ'' zwei reguläre Stellen von $\mathfrak{A}$, so ist

$$(\lambda' - \lambda'')\,\mathfrak{R}_{\lambda'}\,\mathfrak{R}_{\lambda''} = \mathfrak{R}_{\lambda''} - \mathfrak{R}_{\lambda'}.$$

Es sei π eine beliebige beschränkte und abgeschlossene Punktmenge auf der λ-Ebene, die wir wieder mit der Φ-Ebene identifizieren können. Den trivialen Fall, wo π nur endlich viele Punkte enthält, können wir ausschließen und daher auf π nach bekannten Sätzen eine (nötigenfalls auf π überall dichte) Punktfolge $\pi' = \{\lambda_1, \lambda_2, \ldots\}$ derart bestimmen, daß π mit der abgeschlossenen Hülle von π' identisch ist, also aus den Punkten λ_i sowie ihren Häufungspunkten besteht. Bezeichnet $\mathsf{R}(\mathfrak{A})$ die nach Voraussetzung endliche obere Grenze aller $|\lambda_i|$, so ist die Kopplungsform der Diagonalmatrix $\mathfrak{A} = \|\lambda_i \delta_{ik}\|$ auf E absolut gewiß $\leqq \mathsf{R}(\mathfrak{A})$, also ist $\mathfrak{A}$ beschränkt. Man überzeugt sich leicht, daß die zu $\lambda\,\mathfrak{E} - \mathfrak{A}$ gehörigen inhomogenen Gleichungen $(\lambda - \lambda_i)x_i = c_i$ für jede quadratisch konvergente Zahlenfolge $\{c_i\}$ dann und nur dann genau eine quadratisch konvergente Lösung haben, wenn λ weder ein λ_i noch ein Häufungspunkt der λ_i ist. M. a. W., die singulären Stellen von $\mathfrak{A}$ sind mit den Punkten der beschränkten und abgeschlossenen Punktmenge π identisch, die willkürlich vorgeschrieben werden kann.

Wir zeigen jetzt umgekehrt, daß die Menge der singulären Stellen bei jeder beschränkten Matrix $\mathfrak{A}$ beschränkt und abgeschlossen ist. Zunächst müssen sie alle im Kreisbereiche $|\Phi| \leqq \mathsf{P}(\mathfrak{A}) < +\infty$ liegen. Wegen (231) ist nämlich $\mathsf{P}(\mathfrak{A}^m) \leqq (\mathsf{P}(\mathfrak{A}))^m$, $m = 1, 2, \ldots$, so daß die C. Neumannsche Reihe $\mathfrak{P}_\lambda = \sum\limits_0^\infty \frac{\mathfrak{A}^m}{\lambda^{m+1}}$ mit Rücksicht auf die bei (230) gemachte Bemerkung für $|\lambda| > \mathsf{P}(\mathfrak{A})$ gewiß konvergiert und eine beschränkte Matrix darstellt. Man verifiziert leicht, wie auf S. 136, daß sie sowohl vordere als auch hintere Reziproke von $\lambda\,\mathfrak{E} - \mathfrak{A}$ ist, w. z. b. w.

[Aus der Verifikation erkennt man, daß, sobald $\mathfrak{P}_\lambda$ etwa für $|\lambda| > \varrho$ konvergiert und eine beschränkte Matrix darstellt, für $|\lambda| > \varrho$ kein λ singulär ist, so daß $\mathfrak{x} = \mathfrak{P}_\lambda\,\mathfrak{c}$ die einzige quadratisch konvergente Lösung von $(\lambda\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{x} = \mathfrak{c}$, $|\mathfrak{c}|^2 < +\infty$ ist. — Man kann den Beschränktheitsbegriff auf nichtlineare Gleichungssysteme (mit unendlichen Potenzreihen unendlich vieler Variablen) restlos verallgemeinern und die soeben gemachte Bemerkung leistet dabei beim Existenzbeweis der

Lösung oder bei analytischen Fortsetzungen der Lösung einen wesentlichen Dienst.]

Wir haben noch zu zeigen, daß die Menge der singulären Punkte immer abgeschlossen ist. Zu diesem Ende genügt es zu beweisen, daß jede reguläre Stelle in einem Gebiet von lauter regulären Stellen liegt. Ist aber $\mathsf{n}(\mathfrak{A}\mathfrak{A}^*) > 0$, so gibt es, wie man sich leicht überzeugt, aus Gründen der Stetigkeit ein $\varepsilon > 0$ derart, daß im Kreise $|\lambda| < \varepsilon$ auch $\mathsf{n}((\lambda\mathfrak{E} - \mathfrak{A})(\lambda\mathfrak{E} - \mathfrak{A})^*) > 0$ gilt und entsprechendes kann für die vordere Norm gezeigt werden. Wir können aber auch mehr behaupten: ist $\lambda = \lambda_0$ eine (im Endlichen gelegene) reguläre Stelle, so daß die eindeutige Reziproke $(\lambda_0\mathfrak{E} - \mathfrak{A})^{-1}$ existiert, so kann diese nach aufsteigenden ganzen Potenzen von $\lambda - \lambda_0$ in eine Matrizenpotenzreihe entwickelt werden, die für $|\lambda - \lambda_0| < \varepsilon$ eine konvergente Darstellung einer beschränkten, zu $\lambda\mathfrak{E} - \mathfrak{A}$ von vorn und von hinten reziproken Matrix ist. Durch eine Verlegung des Nullpunktes können wir erreichen, daß $\lambda_0 = 0$ wird. Dann ist also $\mathfrak{A}^{-1}$ vorhanden (mit $\mathfrak{A}^{-1}$ wollen wir nur eine Reziproke bezeichnen, für welche der Fall 1 vorliegt). Die Reihe $-\sum_{n=0}^{\infty} \lambda^n (\mathfrak{A}^{-1})^{n+1}$ ist dann für $|\lambda| < \mathsf{P}(\mathfrak{A}^{-1})$ konvergent und stellt, wie eigentlich bereits auf S. 136 verifiziert wurde, eine sowohl vordere als auch hintere beschränkte Reziproke von $\lambda\mathfrak{E} - \mathfrak{A}$ dar, w. z. b. w. Alle diese Entwicklungen entstehen aus der Neumannschen, die zu $\lambda_0 = \infty$ gehört, durch eine projektive Substitution. — Da in der Umgebung jeder regulären Stelle λ_0 die Matrix $\|R_{ik}(\lambda)\| = (\lambda\mathfrak{E} - \mathfrak{A})^{-1}$ nach aufsteigenden Potenzen von $\lambda - \lambda_0$ entwickelt werden kann, so gilt dies erst recht von den Funktionen $R_{ik}(\lambda)$. Wir denken uns dabei aus der Ebene die singulären Stellen entfernt und dadurch die $R_{ik}(\lambda)$ evtl. künstlich eindeutig gemacht, damit es sich um den Zweig von $R_{ik}(\lambda)$ handeln soll, der einen unmittelbaren matrizentheoretischen Sinn hat. Es ist unnötig zu betonen, daß das Gebiet der regulären Stellen von $\mathfrak{A}$ nicht einfach zusammenhängend, ja nicht einmal zusammenhängend zu sein braucht, so daß $R_{ik}(\lambda)$ in verschiedenen Gebieten der Ebene verschiedenen analytischen Funktionen oder „verschiedenen" Zweigen derselben analytischen Funktion angehören kann.

In einer von i und k unabhängigen Umgebung jeder matrizentheoretisch regulären Stelle von $\mathfrak{A}$ ist, wie wir soeben gesehen haben, jede der Funktionen $R_{ik}(\lambda)$ gewiß regulär. Daraus folgt, daß jede beschränkte Matrix $\mathfrak{A}$ mindestens eine singuläre Stelle hat. Denn sonst wäre jedes $R_{ik}(\lambda)$ eine ganze Funktion, die mit Rücksicht auf die in der Umgebung der Stelle $\lambda_0 = \infty$ gültigen Neumannschen Entwicklung für $|\lambda| \to +\infty$ gleichmäßig verschwindet, so daß nach dem

Satze von Liouville $R_{ik}(\lambda) \equiv 0$ ist. Dies ist aber eine Absurdität, da $\| R_{ik}(\lambda) \|$ mit $(\lambda\mathfrak{E} - \mathfrak{A})$ multipliziert mindestens für $|\lambda| > \mathsf{P}(\mathfrak{A})$ die Einheitsmatrix ergeben soll, also nicht die Nullmatrix sein kann. — Implizite haben wir bewiesen, daß mindestens ein $R_{ik}(\lambda)$ mindestens eine singuläre Stelle hat.

Der Satz, daß die Menge der matrizentheoretisch singularen Stellen von $\mathfrak{A}$ aus den funktionentheoretischen Singularitäten der Resolvente und *nur* aus diesen Stellen besteht, ist wohl bei endlichen Matrizen (S. 17), nicht aber bei den beschränkten richtig. Das Beispiel (242) belegt dies. Die Menge der singulären Stellen dieser Matrix ist nämlich, wie sich leicht zeigen läßt, identisch mit dem Kreisbereiche $|\lambda| \leqq 1$, trotzdem jedes $R_{ik}(\lambda)$ mit Ausnahme des einzigen Punktes $\lambda = 0$ durchweg regulär ist. Vgl. S. 152. Allerdings ist (242) nicht Hermitesch. — Vgl. S. 178 weiter unten.

§ 64. Die Hermiteschen Abschnittsspektra.

Es sei $\mathfrak{A}$ eine Hermitesche Matrix und es sei K ein (offenes) Kreisgebiet derart, daß es keinen Punkt des Abschnittspektrums von $\mathfrak{A}$ enthält, d. h. es möge jeder Eigenwert eines jeden Abschnittes $\mathfrak{A}_{[n]}$ von $\mathfrak{A}$ außerhalb K liegen. — Solche Kreise gibt es gewiß, da jeder Abschnittseigenwert absolut $\leqq \mathsf{P}(\mathfrak{A}_{[n]}) \leqq \mathsf{P}(\mathfrak{A}) < +\infty$ ist. Wir werden später (§ 102) auch für nicht beschränkte Matrizen einen Satz beweisen, aus welchem folgt, daß jeder Punkt von K regulärer Punkt für $\mathfrak{A}$ ist; genauer, bezeichnet $\| R_{ik}^{(n)}(\lambda) \|_{(n)}$ die (auf K gewiß vorhandene) Resolvente von $\mathfrak{A}_{[n]}$, so ist auf K der Grenzwert $\lim\limits_{n=\infty} R_{ik}^{(n)}(\lambda)$ bei jedem i und k vorhanden und die Matrix $\| \lim\limits_{n=\infty} R_{ik}^{(n)}(\lambda) \|$ stellt dabei eine, sowohl hintere als auch vordere, beschränkte Reziproke von $\lambda\mathfrak{E} - \mathfrak{A}$ dar. Man könnte meinen, die dadurch gelieferte Bedingung sei nicht nur hinreichend, sondern auch notwendig dafür, daß λ^0 keine singuläre Stelle von $\mathfrak{A}$ ist. Das Toeplitzsche Beispiel der Matrix

$$(246) \qquad \mathfrak{A} = \left\| \begin{matrix} 0 & 1 & & & \\ 1 & 0 & & & \\ & & 0 & 1 & \\ & & 1 & 0 & \\ & & & & \ddots \end{matrix} \right\|,$$

die durch Verschiebung „irreduzibler Quadrate“ $\left\| \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\|$ entlang der

Diagonale entsteht und sonst mit lauter Nullen besetzt ist, belehrt uns jedoch des Gegenteils[1]). Zunächst ist, wie man sich leicht überzeugt, $\mathfrak{A}^2 = \mathfrak{E}$. Also ist $\mathfrak{A}$ Reziproke von sich selbst, mithin hat $\lambda\,\mathfrak{E} - \mathfrak{A}$ für $\lambda = 0$ eine vordere und hintere Reziproke, nämlich $-\mathfrak{A}$, die offenbar beschränkt ist. Also ist $\lambda = 0$ eine reguläre Stelle von $\mathfrak{A}$. Trotzdem gibt es unendlich viele verschwindende Abschnittseigenwerte. Denn in $\mathfrak{A}_{[2n+1]}$ wird die letzte Zeile von lauter Nullen gebildet, so daß $\det \mathfrak{A}_{[2n+1]} = 0$ ist und daher ein verschwindender $(2\,n + 1)$-ter Eigenwert existieren muß. — Indem man etwa die getrennten irreduziblen Quadrate binär auf die Hauptachsen transformiert, kann man hier übrigens auch das ganze Abschnittspektrum ablesen: $\mathfrak{A}_{[n]}$ hat $\left[\frac{n}{2}\right]$ Eigenwerte $= +1$ und $\left[\frac{n}{2}\right]$ Eigenwerte $= -1$; hierbei ist, wie üblich, $n = 2\,m + 1$ oder $n = 2\,m$ und beidemal $\left[\frac{n}{2}\right] = m$ gesetzt. Es ließe sich ohne Mühe zeigen, daß die beiden Punkte $\lambda = \pm 1$, im Gegensatz zu $\lambda = 0$, bereits singuläre Stellen von $\mathfrak{A}$ sind. Dies ist kein Zufall, es folgt vielmehr auch aus einem allgemeinen Satz über Hermitesche Matrizen, denen wir uns jetzt im besonderen zuwenden.

§ 65. Ein Satz über Hermitesche Matrizen.

Ist $\mathfrak{A}$ von Hermiteschem Typus, so ist $\lambda\,\mathfrak{E} - \mathfrak{A}$, wenn λ nicht reell ist, zwar nicht von Hermiteschem Typus, jedoch, wegen $\mathfrak{A} = \mathfrak{A}^*$ und da mit $\lambda\,\mathfrak{E}$ jede Matrix vertauschbar ist, offenbar normal, so daß für $\lambda\,\mathfrak{E} - \mathfrak{A}$ der Fall **3** gewiß ausgeschlossen ist. Um also die regulären Stellen von $\mathfrak{A}$ aufzufinden, brauchen wir nur etwa die hintere Norm von $\lambda\,\mathfrak{E} - \mathfrak{A}$ zu betrachten. Die Bedingung **1** in (241) ergibt, daß λ für $\mathfrak{A}$ dann und nur dann regulär ist, wenn $\mathsf{n}((\lambda\,\mathfrak{E} - \mathfrak{A})(\lambda\,\mathfrak{E} - \mathfrak{A})^*) > 0$ ausfällt. Daraus folgt, daß alle nicht reellen λ reguläre Stellen sind. Es sei λ reell. Dann ist $(\lambda - \mathfrak{A})^* = \bar{\lambda} - \mathfrak{A}^* = \lambda - \mathfrak{A}$, und unsere Ungleichheit geht in $\mathsf{n}((\lambda\,\mathfrak{E} - \mathfrak{A})^2) > 0$ über. Diese Bedingung ist, wie man sich leicht überzeugt, für $\lambda < -\mathsf{P}(\mathfrak{A})$ sowie für $\lambda > \mathsf{P}(\mathfrak{A})$ gewiß erfüllt, doch brauchen wir dies nicht besonders zu beweisen, da für $|\lambda| > \mathsf{P}(\mathfrak{A})$ die Neumannschen Reihen immer eine beschränkte Reziproke liefern. Auf dem Intervall $-\mathsf{P}(\mathfrak{A}) \leqq \lambda \leqq \mathsf{P}(\mathfrak{A})$ muß es mindestens eine singuläre

[1]) Umgekehrt, es gibt beschränkte (nicht Hermitesche) Matrizen $\mathfrak{A}$ derart, daß eine Stelle λ zu den matrizentheoretisch singulären Stellen von $\mathfrak{A}$ gehört und dennoch keine Häufungsstelle von Abschnittseigenwerten ist. Des näheren vgl. S. 152.

Stelle geben, da sonst keine singuläre Stelle vorhanden sein würde. Mit Rücksicht auf (213) behaupten wir mehr, wenn wir sagen, daß bereits auf dem Intervall $\mathsf{n}(\mathfrak{A}) \leqq \lambda \leqq \mathsf{m}(\mathfrak{A}) \left(\gtreqless 0\right)$ mindestens eine singuläre Stelle liegen muß. Dies folgt aber daraus, daß für $\lambda < \mathsf{n}(\mathfrak{A})$ und für $\lambda > \mathsf{m}(\mathfrak{A})$ jedes λ regulär ist. Denn $\mathsf{n}(\mathfrak{A})$ ist nach Definition (S. 124) die untere, $\mathsf{m}(\mathfrak{A})$ die obere Grenze aller Abschnittseigenwerte, so daß auf den beiden Halbgeraden $\lambda < \mathsf{n}(\mathfrak{A})$, $\lambda > \mathsf{m}(\mathfrak{A})$ kein einziger Abschnittseigenwert liegt.

Wir zeigen nun, daß die *Endpunkte* des Abschnittspektrums, ${}_0\lambda = \mathsf{n}(\mathfrak{A})$ und ${}^0\lambda = \mathsf{m}(\mathfrak{A})$, die im Falle $\mathfrak{A} = c\,\mathfrak{E}$, $c \gtreqless 0$ zusammenfallen, gewiß singuläre Stellen von $\mathfrak{A}$ sind. Die Behauptung besagt, daß $\lambda = {}^0\lambda$ und $\lambda = {}^0\lambda$ Wurzeln der Gleichung $\mathsf{n}\left((\lambda\,\mathfrak{E} - \mathfrak{A})^2\right) = 0$ sind. Wir können uns dabei auf den Fall des oberen Endpunktes ${}^0\lambda$ beschränken. Denn sonst würde man von $\mathfrak{A}$ zu $-\mathfrak{A}$ übergehen. Wir können ferner annehmen, daß der untere Endpunkt ${}_0\lambda$ in den Nullpunkt $\lambda = 0$ hineinfällt. Denn sonst würde man $\mathfrak{A}$ durch $\mathfrak{A} - {}_0\lambda\,\mathfrak{E}$ ersetzen können, weil ja $\Phi({}_0\lambda\,\mathfrak{E}; \mathfrak{x}) \equiv {}_0\lambda$ ist für $|\mathfrak{x}| = 1$. Wir haben also nur zu zeigen, daß $\mathsf{n}\left(({}^0\lambda\,\mathfrak{E} - \mathfrak{A})^2\right) = 0$ wird, wenn

(247) $$ {}^0\lambda = \mathsf{m}(\mathfrak{A}) \geqq \mathsf{n}(\mathfrak{A}) = {}_0\lambda = 0 $$

gilt; $\mathfrak{A}$ ist nichtnegativ definit. Aus (213), (247) folgt zunächst $\mathsf{m}(\mathfrak{A}) = \mathsf{M}(\mathfrak{A}) = \mathsf{P}(\mathfrak{A}) = {}^0\lambda$, und daraus wegen (235) und mit Rücksicht auf den Hermiteschen Charakter $\mathfrak{A} = \mathfrak{A}^*$ offenbar

(248) $$ {}^0\lambda = \mathsf{M}(\mathfrak{A}), \qquad \mathsf{M}(\mathfrak{A}^2) = (\mathsf{M}(\mathfrak{A}))^2. $$

Nun beachte man, daß nach der Schwarzschen Ungleichung für $|\mathfrak{x}| = 1$

(249) $$ [\Phi(\mathfrak{A};\mathfrak{x})]^2 = \Big[\sum_{i=1}^{\infty}\Big(\sum_{k=1}^{\infty} a_{ik}x_k\Big)\bar{x}_i\Big]^2 \leqq \sum_{i=1}^{\infty}\Big|\sum_{k=1}^{\infty} a_{ik}x_k\Big|^2 = \Phi(\mathfrak{A}^2;\mathfrak{x}) \leqq \mathsf{M}(\mathfrak{A}^2) $$

gilt. Vgl. nämlich (233), wobei $\mathfrak{A}^* = \mathfrak{A}$ zu setzen ist. Aus (248) und (249) folgt für jeden normierten Vektor $\mathfrak{x}$

(250) $$ [\Phi(\mathfrak{A};\mathfrak{x})]^2 \leqq \Phi(\mathfrak{A}^2;\mathfrak{x}) \leqq {}^0\lambda^2. $$

Wegen ${}^0\lambda = \mathsf{M}(\mathfrak{A}) = \mathsf{m}(\mathfrak{A})$ gibt es ferner eine Folge $\{\mathfrak{x}_n\}$ von normierten Vektoren derart, daß $\Phi(\mathfrak{A};\mathfrak{x}_n) \to {}^0\lambda$ gilt für $n \to +\infty$. Wegen (250) muß für dieselbe Folge auch $\Phi(\mathfrak{A}^2;\mathfrak{x}_n) \to {}^0\lambda^2$, also wegen $\Phi(\mathfrak{E};\mathfrak{x}_n) = |\mathfrak{x}_n|^2 = 1$ offenbar

(251) $$ \Phi(\mathfrak{A};\mathfrak{x}_n) \to {}^0\lambda, \qquad \Phi(\mathfrak{A}^2;\mathfrak{x}_n) \to {}^0\lambda^2, \qquad \Phi(\mathfrak{E};\mathfrak{x}_n) \to 1 $$

gelten. Aus (251) folgt nun

$$\Phi(({}^0\lambda\,\mathfrak{E}-\mathfrak{A})^2;\mathfrak{x}_n)\equiv{}^0\lambda^2\,\Phi(\mathfrak{E};\mathfrak{x}_n)-2\,{}^0\lambda\,\Phi(\mathfrak{A};\mathfrak{x}_n)$$
$$+\,\Phi(\mathfrak{A}^2;\mathfrak{x}_n)\to{}^0\lambda^2\cdot 1-2\,{}^0\lambda\cdot{}^0\lambda+{}^0\lambda^2=0,$$

so daß die Kopplungsform der nichtnegativ definiten Hermiteschen Matrix $({}^0\lambda\,\mathfrak{E}-\mathfrak{A})^2$ unter der Nebenbedingung $|\mathfrak{x}|=1$ der Null beliebig nahe kommen kann. M. a. W., es ist $\mathsf{n}(({}^0\lambda\,\mathfrak{E}-\mathfrak{A})^2)=0$, w. z. b. w.

Da die reellen Diagonalmatrizen von Hermiteschem Typus sind, so erkennt man, wie auf S. 143, daß jede auf dem (evtl. auf einen Punkt zusammenschrumpfenden) Intervall $\mathsf{n}(\mathfrak{A})\leqq\lambda\leqq\mathsf{m}(\mathfrak{A})$ gelegene abgeschlossene Punktmenge, also z. B. dieses ganze Intervall oder darauf getrennt liegende Intervalle die singuläre Menge einer passend zu wählenden beschränkenden Hermiteschen Matrix bilden können (ebenso wie bei beliebigen beschränkten Matrizen die singuläre Menge z. B. aus Kreisscheiben, aber auch aus Kreisringen usf. bestehen kann). Auf die spektrale Analyse dieser Punktmenge kommen wir noch ausführlich zurück.

§ 66. Über die Lage der matrizentheoretisch singulären Stellen.

Es bezeichne $\mathsf{S}(\mathfrak{A})$ die Menge der matrizentheoretisch singulären Stellen der beschränkten und nicht notwendig Hermiteschen Matrix $\mathfrak{A}$. Nach § 63 ist $\mathsf{S}(\mathfrak{A})$ eine nicht leere, abgeschlossene und beschränkte Punktmenge der komplexen λ-Ebene. Die Zahl λ liegt dann und nur dann in $\mathsf{S}(\mathfrak{A})$, wenn $(\lambda\,\mathfrak{E}-\mathfrak{A})^{-1}$ nicht existiert. Hierfür ist nach (241) notwendig und hinreichend, daß mindestens eine der beiden reellen und nichtnegativen Zahlen

$$\mathsf{n}((\lambda\,\mathfrak{E}-\mathfrak{A})(\lambda\,\mathfrak{E}-\mathfrak{A})^*),\qquad \mathsf{n}((\lambda\,\mathfrak{E}-\mathfrak{A})^*(\lambda\,\mathfrak{E}-\mathfrak{A}))$$

verschwindet. Es ist also wegen

$$|\Phi(\mathfrak{C};\mathfrak{x})|^2\leqq\Phi(\mathfrak{C}\,\mathfrak{C}^*;\mathfrak{x}),\qquad |\Phi(\mathfrak{C};\mathfrak{x})|^2\leqq\Phi(\mathfrak{C}^*\,\mathfrak{C};\mathfrak{x})$$
$$(|\mathfrak{x}|=1;\ \mathfrak{C}=\lambda\,\mathfrak{E}-\mathfrak{A})$$

und mit Rücksicht auf

$$\Phi(\mathfrak{C};\mathfrak{x})=\Phi(\lambda\,\mathfrak{E}-\mathfrak{A};\mathfrak{x})=\lambda\,\Phi(\mathfrak{E};\mathfrak{x})-\Phi(\mathfrak{A};\mathfrak{x})=\lambda-\Phi(\mathfrak{A};\mathfrak{x})$$

offenbar notwendig, daß

$$\lambda-\Phi(\mathfrak{A};\mathfrak{x})$$

unter der Nebenbedingung $|\mathfrak{x}| = 1$ der Null beliebig nahe kommen kann, d. h. daß die Zahl λ in dem Wertbereiche $\mathsf{W}(\mathfrak{A})$ der Kopplungsform liegt. Mithin ist jeder Punkt von $\mathsf{S}(\mathfrak{A})$ auch in $\mathsf{W}(\mathfrak{A})$ enthalten.

Für endliche Matrizen haben wir diesen Satz bereits in § 19 bewiesen. Man könnte also glauben, daß die soeben vorgenommene Übertragung auf unendliche Matrizen nichts Neues aussagt. In der Tat haben wir vorher dem Toeplitzschen Satze (§ 19) unmittelbar entnehmen können, daß das Abschnittsspektrum der unendlichen Matrix $\mathfrak{A}$ stets in $\mathsf{W}(\mathfrak{A})$ liegt. Nun gibt es aber, wie bereits erwähnt (S. 146), beschränkte (nicht Hermitesche) Matrizen $\mathfrak{A}$, bei welchen $\mathsf{S}(\mathfrak{A})$ solche Punkte enthält, die weder Abschnittseigenwerte noch deren Häufungswerte sind, so daß die obige direkte Schlußweise ebenso unumgänglich war, wie in § 63 der Existenzbeweis für $\mathsf{S}(\mathfrak{A})$. Zu jeder beschränkten und abgeschlossenen Punktmenge Π gibt es bekanntlich einen und nur einen konvexen Bereich (vgl. S. 34), der unter allen, Π enthaltenden konvexen Bereichen der *kleinste* ist; er heißt konvexe Hülle von Π. Es bezeichne $\mathsf{T}(\mathfrak{A})$ die konvexe Hülle von $\mathsf{S}(\mathfrak{A})$. Da $\mathsf{S}(\mathfrak{A})$ in $\mathsf{W}(\mathfrak{A})$ liegt und $\mathsf{W}(\mathfrak{A})$ nach § 18 konvex ist, so liegt offenbar auch $\mathsf{T}(\mathfrak{A})$ in $\mathsf{W}(\mathfrak{A})$, ohne notwendig mit $\mathsf{W}(\mathfrak{A})$ identisch zu sein. Ist aber $\mathsf{P}(\mathfrak{A}) = \mathsf{M}(\mathfrak{A})$, so muß $\mathsf{T}(\mathfrak{A}) = \mathsf{W}(\mathfrak{A})$ sein. In dem Spezialfall, wo $\mathfrak{A}$ Hermitesch ist [also nach (213) der Bedingung $\mathsf{P}(\mathfrak{A}) = \mathsf{M}(\mathfrak{A})$ gewiß genügt], ist hierfür der Beweis in § 65 geführt worden. In dem allgemeinen Falle ist der Beweis etwas komplizierter, verläuft aber sonst ganz analog, so daß ich darauf hier nicht eingehen will. Es sei nur erwähnt, daß es sich um eine Verallgemeinerung der sich auf endliche und normale Matrizen beziehenden Toeplitzschen Polygonregel (§ 19) handelt; denn in diesem Falle ist die Bedingung $\mathsf{P} = \mathsf{M}$ nach § 20 gewiß erfüllt. —

Aus den Formalsätzen von § 59 folgt ebenso wie bei den endlichen Matrizen (S. 7), daß die vier beschränkten Matrizen $\mathfrak{B}, \mathfrak{B}', \overline{\mathfrak{B}}, \mathfrak{B}^*$ nur zugleich eine eindeutige und beschränkte Reziproke zulassen können, und daß

$$(\mathfrak{B}')^{-1} = (\mathfrak{B}^{-1})', \qquad (\overline{\mathfrak{B}})^{-1} = \overline{(\mathfrak{B}^{-1})}, \qquad \text{also} \qquad (\mathfrak{B}^*)^{-1} = (\mathfrak{B}^{-1})^*$$

gilt. Setzt man $\mathfrak{B} = \lambda\,\mathfrak{E} - \mathfrak{A}$, so folgt daraus wegen

$$(\lambda\,\mathfrak{E} - \mathfrak{A})' = \lambda\,\mathfrak{E} - \mathfrak{A}', \quad \overline{(\lambda\,\mathfrak{E} - \mathfrak{A})} = \bar{\lambda}\,\mathfrak{E} - \overline{\mathfrak{A}}, \quad (\lambda\,\mathfrak{E} - \mathfrak{A})^* = \bar{\lambda}\,\mathfrak{E} - \mathfrak{A}^*$$

offenbar

$$\mathsf{S}(\mathfrak{A}') = \mathsf{S}(\mathfrak{A}), \quad \mathsf{S}(\overline{\mathfrak{A}}) = \overline{\mathsf{S}(\mathfrak{A})}, \quad \text{also} \quad \mathsf{S}(\mathfrak{A}^*) = \overline{\mathsf{S}(\mathfrak{A})},$$

wobei $\overline{\mathsf{S}(\mathfrak{A})}$ die Punktmenge bezeichnet, in welche $\mathsf{S}(\mathfrak{A})$ durch Spiegelung an der reellen Achse übergeht. Diese Beziehungen, die im

Falle endlicher Matrizen auch mittels Determinanten verifiziert werden können (S. 23), stehen im Einklang damit, daß

$$\mathsf{W}(\mathfrak{A}') = \mathsf{W}(\mathfrak{A}), \quad \mathsf{W}(\overline{\mathfrak{A}}) = \overline{\mathsf{W}(\mathfrak{A})}, \quad \text{also} \quad \mathsf{W}(\mathfrak{A}^*) = \overline{\mathsf{W}(\mathfrak{A})}$$

ist. — Es kann endlich, etwa mittels den C. Neumann-Hilbschen Reihen, ohne Mühe gezeigt werden, daß, wenn $\mathfrak{A}^{-1}$ existiert, d. h. wenn der Punkt $\lambda = 0$ außerhalb $\mathsf{S}(\mathfrak{A})$ liegt, auch jetzt (vgl. S. 15)

$$\mathsf{S}(\mathfrak{A}^{-1}) = \frac{1}{\mathsf{S}(\mathfrak{A})}$$

gilt, wobei $\frac{1}{\mathsf{S}(\mathfrak{A})}$ die Menge der reziproken Werte der in $\mathsf{S}(\mathfrak{A})$ enthaltenen Zahlen bedeutet.

§ 67. Die Resolvente in dem pathologischen Falle.

Ist $\mathfrak{S}$ eine beschränkte Matrix, für welche $\mathfrak{S}^{-1}$ existiert, so wollen wir wieder sagen, die Matrix $\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$ entstehe aus der beschränkten Matrix $\mathfrak{A}$ durch Transformation mittels $\mathfrak{S}$. Die transformierte Matrix ist nach § 55 wieder beschränkt. Wir wollen beweisen, daß die Menge der matrizentheoretisch singulären Stellen Matrizentransformationen gegenüber invariant ist. Es bezeichne zu diesem Ende λ eine außerhalb $\mathsf{S}(\mathfrak{A})$ gelegene Zahl, so daß $(\lambda\mathfrak{E} - \mathfrak{A})^{-1}$ existiert. Dann gilt wegen

$$\mathfrak{S}(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{S}^{-1} = \lambda\mathfrak{E} - \mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$$

und mit Rücksicht auf

$$\mathfrak{S}(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{S}^{-1}\mathfrak{S}(\lambda\mathfrak{E} - \mathfrak{A})^{-1}\mathfrak{S}^{-1} = \mathfrak{E}$$

offenbar

$$(\lambda\mathfrak{E} - \mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})(\mathfrak{S}(\lambda\mathfrak{E} - \mathfrak{A})^{-1}\mathfrak{S}^{-1}) = \mathfrak{E},$$

und ebenso folgt

$$(\mathfrak{S}(\lambda\mathfrak{E} - \mathfrak{A})^{-1}\mathfrak{S}^{-1})(\lambda\mathfrak{E} - \mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}) = \mathfrak{E},$$

so daß $\mathfrak{S}(\lambda\mathfrak{E} - \mathfrak{A})^{-1}\mathfrak{S}^{-1}$ eine vordere und zugleich hintere und dabei nach dem ersten Faltungssatz beschränkte Reziproke von $\lambda\mathfrak{E} - \mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$ ist. Folglich liegt jede außerhalb $\mathsf{S}(\mathfrak{A})$ gelegene Zahl außerhalb $\mathsf{S}(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})$. Da $\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}$ durch $\mathfrak{S}^{-1}$ in $\mathfrak{A}$ transformiert wird, so folgt daraus, daß auch jede außerhalb $\mathsf{S}(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1})$ gelegene Zahl außerhalb $\mathsf{S}(\mathfrak{A})$ liegen muß. Mithin ist

$$\mathsf{S}(\mathfrak{A}) = \mathsf{S}(\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1}),$$

w. z. b. w.

Es sei nun $\mathfrak{A}$ derart, daß $\mathfrak{A}^{-1}$ existiert. Dann ist

$$\mathfrak{A}^* \mathfrak{A} = \mathfrak{A}^{-1} \mathfrak{A} \mathfrak{A}^* \mathfrak{A},$$

d. h. die beiden Normmatrizen gehen durch eine Matrizentransformation ($\mathfrak{S} = \mathfrak{A}^{-1}$) ineinander über, so daß sie mit Rücksicht auf die soeben bewiesene Relation dieselbe singuläre Punktmenge haben:

$$(*) \qquad \mathsf{S}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{S}(\mathfrak{A}^*\mathfrak{A}).$$

Dieser Satz ist, im Gegensatz zu den endlichen Matrizen (§ 14), nicht allgemein richtig, wenn $\mathfrak{A}^{-1}$ nicht existiert. Denn z. B. die eine der beiden Normmatrizen von (245) ist die Einheitsmatrix, die andere aber diejenige Diagonalmatrix, welche aus der Einheitmatrix dadurch entsteht, daß das $(1, 1)$-te Element $(= 1)$ durch Null ersetzt wird. Die Menge der matrizentheoretisch singulären Stellen enthält also bei der einen Normmatrix nur den Punkt $\lambda = 1$, bei der anderen aber auch den Punkt $\lambda = 0$, so daß $\mathsf{S}(\mathfrak{A}\mathfrak{A}^*) \neq \mathsf{S}(\mathfrak{A}^*\mathfrak{A})$ wird.

Es sei jetzt wieder vorausgesetzt, daß $\mathfrak{A}^{-1}$ existiert, und daher (*) gewiß richtig ist. Die beiden Normmatrizen $\mathfrak{A}\mathfrak{A}^*$, $\mathfrak{A}^*\mathfrak{A}$ sind auch bei nicht Hermiteschem $\mathfrak{A}$ Hermitesch und nichtnegativ definit. Aus (*) folgt daher nach § 65

$$(**) \qquad \mathsf{n}(\mathfrak{A}\mathfrak{A}^*) = \mathsf{n}(\mathfrak{A}^*\mathfrak{A}),$$

wenn $\mathfrak{A}^{-1}$ existiert. Dies gilt, wenn nur $\mathfrak{A}^{-1}$ existiert, so daß die Toeplitzsche Tabelle (241) dahin verschärft werden kann, daß im Falle **1** die beiden Zahlen (**) gleich sein müssen. Anders ausgedrückt: sie sind stets nichtnegativ, können auch verschieden sein, sind aber gewiß gleich, wenn sie positiv sind.

Wir können einen Schritt noch weiter gehen. Die Voraussetzung, daß $\mathfrak{A}^{-1}$ existiert, d. h. daß der Punkt $\lambda = 0$ außerhalb $\mathsf{S}(\mathfrak{A})$ liegt, lassen wir jetzt nämlich fallen und machen nur die Annahme, daß es gegen $\lambda = 0$ konvergierende Zahlen $\lambda_1, \lambda_2, \ldots$ gibt, die außerhalb $\mathsf{S}(\mathfrak{A})$ liegen [diese Annahme ist tatsächlich weniger einschränkend als die frühere; denn $\mathsf{S}(\mathfrak{A})$ ist eine abgeschlossene Punktmenge]. Aus (**) folgt dann, da $(\lambda_m \mathfrak{E} - \mathfrak{A})^{-1}$ für $m = 1, 2, \ldots$ existiert,

$$\mathsf{n}((\lambda_m \mathfrak{E} - \mathfrak{A})(\lambda_m \mathfrak{E} - \mathfrak{A})^*) = \mathsf{n}((\lambda_m \mathfrak{E} - \mathfrak{A})^*(\lambda_m \mathfrak{E} - \mathfrak{A}))$$

für jedes m. Wegen $\lim\limits_{m=\infty} \lambda_m = 0$ gilt daher aus leicht ersichtlichen Stetigkeitsgründen auch (**) selbst. Nur werden jetzt beide Zahlen (**) gleich Null sein können (wenn nämlich $\mathfrak{A}^{-1}$ nicht existiert).

Wir haben damit den folgenden Satz bewiesen: Die pathologischen Fälle 3 (S. 138) können nur bei solchen Matrizen $\mathfrak{A}$ vorliegen, bei welchen es ein $\varepsilon > 0$ gibt derart, daß jeder Punkt des Kreisbereiches $|\lambda| \leqq \varepsilon$ der matrizentheoretisch singulären Punktmenge $\mathsf{S}(\mathfrak{A})$ angehört (übrigens ist diese notwendige Bedingung noch weitaus nicht hinreichend).

Dieser Sachverhalt kann bei der Matrix (245) ohne Mühe verifiziert werden. Die zu $\lambda\mathfrak{E} - \mathfrak{A}$ gehörigen inhomogenen Gleichungen $(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{c}$ sind nämlich

$$\lambda x_1 = c_1, \quad \lambda x_2 - x_1 = c_2, \quad \lambda x_3 - x_2 = c_3, \quad \ldots,$$

woraus durch vollständige Induktion folgt, daß

$$x_i = \sum_{k=1}^{i} \frac{c_k}{\lambda^{i-k+1}}; \qquad i = 1, 2, \ldots$$

ist. Andererseits gilt $\mathfrak{x} = (\lambda\mathfrak{E} - \mathfrak{A})^{-1}\mathfrak{c}$, d. h.

$$x_i = \sum_{k=1}^{\infty} R_{ik}(\lambda)\, c_k; \qquad i = 1, 2, \ldots,$$

wenn wieder $\| R_{ik}(\lambda) \| = (\lambda\mathfrak{E} - \mathfrak{A})^{-1}$ gesetzt wird. Mit Rücksicht auf den rekursiven Charakter des Gleichungssystems ist es offenbar gestattet, die beiden Darstellungen von x_i miteinander auch außerhalb $\mathsf{S}(\mathfrak{A})$ zu identifizieren. So folgt

$$R_{ik}(\lambda) = \frac{1}{\lambda^{i-k+1}} \quad \text{für} \quad k \leqq i \quad \text{und} \quad R_{ik}(\lambda) \equiv 0 \quad \text{für} \quad i < k.$$

Diese Matrix $\| R_{ik}(\lambda) \|$ ist aber nur dann eine Q-Matrix, wenn $|\lambda| > 1$ ist. Erst recht ist diese Bedingung dafür notwendig, daß $\| R_{ik}(\lambda) \|$ beschränkt ist. Folglich gehört der ganze Kreisbereich $|\lambda| \leqq 1$ der matrizentheoretisch singulären Punktmenge $\mathsf{S}(\mathfrak{A})$ an. Andererseits ist die Kopplungsform von (245) unter der Nebenbedingung $|\mathfrak{x}| = 1$ nach der Schwarzschen Ungleichung dem Betrage nach gewiß $\leqq 1$, d. h. $\mathsf{W}(\mathfrak{A})$ liegt in dem Kreise $|\lambda| \leqq 1$. Nach § 66 muß also auch $\mathsf{S}(\mathfrak{A})$ in diesem Kreise liegen. Mithin ist die Menge der matrizentheoretisch singulären Stellen bei der beschränkten Matrix (245) identisch mit dem Kreisbereiche $|\lambda| \leqq 1$.

Das Merkwürdige ist dabei, daß funktionentheoretisch nur der Punkt $\lambda = 0$ singulär ist (d. h. die einzelnen $R_{ik}(\lambda)$ sind mit Ausnahme dieses Punktes durchweg regulär), und daß auch alle Abschnittseigenwerte

von (245) offenbar gleich Null sind. Diese Erscheinungen sind bereits auf S. 145 bzw. S. 146 erwähnt worden.

Wir haben oben auf eine indirekte Weise bewiesen, daß die Abschnittsresolvente der kanonischen Matrix (245) für $|\lambda| > 1$ beschränkt ist. Einfacher folgt dies aus dem folgenden Satze von I. Schur:

Damit die Matrix $\| a_{ik} \|$ beschränkt, ist hinreichend, daß

$$\sum_{k=1}^{\infty} | a_{ik} | < \Omega, \quad i = 1, 2, \ldots; \qquad \sum_{i=1}^{\infty} | a_{ik} | < \Omega, \quad k = 1, 2, \ldots$$

gilt, wobei Ω irgendeine Schranke bedeutet. Dieser Schursche Satz, auf dessen Beweis hier nicht eingegangen werden soll, ist das einzige bisher bekannte allgemein brauchbare hinreichende und dabei explizite Kriterium für die Beschränktheit.

Viertes Kapitel.

Theorie der Spektralmatrix.

§ 68. Unitäre, einseitig unitäre und halbunitäre Matrizen.

Es sei $\mathfrak{U}$ eine Matrix derart, daß $\sum_{j=1}^{\infty} u_{ij}\bar{u}_{kj} = \delta_{ik}$ für alle i und k gilt. Für $i = k$ ist dann die Summe gewiß absolut konvergent, also ist sie es gemäß der Schwarzschen Ungleichung auch für $i \neq k$. Das Gleichungssystem besagt, daß $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$ ist, d. h. daß $\mathfrak{U}^*$ eine hintere Reziproke von $\mathfrak{U}$ darstellt, und zwar ist nach (213), (235)

(252) $(\mathsf{P}(\mathfrak{U}))^2 = (\mathsf{P}(\mathfrak{U}^*))^2 = \mathsf{P}(\mathfrak{U})\,\mathsf{P}(\mathfrak{U}^*) = \mathsf{M}(\mathfrak{U}\mathfrak{U}^*) = \mathsf{M}(\mathfrak{E}) = 1.$

[Hierbei ist stillschweigend vorausgesetzt worden, daß $\mathfrak{U}$ beschränkt ist. Doch braucht dies, wie man sich leicht überzeugt, nicht besonders vorausgesetzt werden: gilt $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$, so ist $\mathfrak{U}$ von selbst beschränkt.]

Auf E ist wegen $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$ die Hermitesche Form $\Phi(\mathfrak{U}\mathfrak{U}^*) = \Phi(\mathfrak{E}) \equiv 1$, also ist $\mathsf{m}(\mathfrak{U}\mathfrak{U}^*) = \mathsf{n}(\mathfrak{U}\mathfrak{U}^*) = 1$, und mithin auch $\mathsf{m}(\mathfrak{U}^*\mathfrak{U}) = 1$, wegen (236). Dies alles folgt aus $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$ allein; eine Matrix $\mathfrak{U}$, für welche diese Bedingung erfüllt ist, nennen wir unitär von rechts. Sie braucht nicht zugleich unitär von links zu sein, d. h. $\mathfrak{U}^*\mathfrak{U}$ ist im allgemeinen $\neq \mathfrak{E}$, m. a. W., $\mathfrak{U}^*$ braucht keine vordere Reziproke von $\mathfrak{U}$ zu sein. Es ist z. B. ersichtlich, daß die unter (242) stehende erste, „Volterrasche" Matrix von rechts unitär ist, während sie nicht auch von links unitär sein kann, da sie unendlich viele hintere, also keine vordere beschränkte Reziproke besitzt, so daß jetzt zu den drei vorher gefundenen Relationen als vierte $\mathsf{n}(\mathfrak{U}^*\mathfrak{U}) = 0$ hinzukommt. Die transponierte Matrix (245) ist freilich von links, aber nicht von rechts unitär; bezeichnet man sie also mit $\mathfrak{U}$, so gilt jetzt umgekehrt $\mathsf{n}(\mathfrak{U}\mathfrak{U}^*) = 0$, $\mathsf{n}(\mathfrak{U}^*\mathfrak{U}) = 1$, $\mathsf{m}(\mathfrak{U}\mathfrak{U}^*) = \mathsf{m}(\mathfrak{U}^*\mathfrak{U}) = 1$. Man erkennt auch einen geometrischen Grund dafür, warum die Möglichkeit, daß aus $\mathfrak{U}\mathfrak{U}^* = \mathfrak{E}$ umgekehrt $\mathfrak{U}^*\mathfrak{U} = \mathfrak{E}$ nicht folgt, wohl bei endlichen (S. 5), nicht aber bei unendlichen Matrizen ausgeschlossen ist, und wieso der Fall 3

jetzt auftreten kann: eine Koordinatentransformation, die, in einen n-dimensionalen Raum eingebettet, eine Drehung ist, braucht nicht bereits für den $(n-1)$-dimensionalen Raum eine („flache") Drehung zu bedeuten; nun gibt es aber zwischen n und $n-1$ nur für $n < +\infty$ einen Unterschied. — Ist $\mathfrak{U}$ von rechts unitär, so gilt, wegen (229) und da $\Phi(\mathfrak{U}\mathfrak{U}^*) = \Phi(\mathfrak{E})$ ist,

$$\sum_{i=1}^{\infty} \left| \sum_{k=1}^{\infty} u_{ik} x_k \right|^2 \equiv \sum_{i=1}^{\infty} x_i \bar{x}_i = 1 \tag{253}$$

auf E. Ist $\mathfrak{U}$ nicht zugleich von links unitär, so gilt auf E zwar $\Phi(\mathfrak{U}^*\mathfrak{U}) \leqq \mathsf{m}(\mathfrak{U}^*\mathfrak{U}) = \mathsf{m}(\mathfrak{U}\mathfrak{U}^*) = 1$, aber es gibt Punkte auf E, wo nicht das Gleichheitszeichen gilt, da sonst auch die untere Grenze $\mathsf{n}(\mathfrak{U}^*\mathfrak{U}) = 1$ wäre, während doch $\mathfrak{U}$ nicht von links unitär sein soll. Ist also $\mathfrak{U}$ von rechts unitär, so ist auf E

$$\sum_{i=1}^{\infty} \left| \sum_{k=1}^{\infty} u_{ki} x_k \right|^2 \leqq 1, \tag{254}$$

wobei das Gleichheitszeichen für *jeden* Punkt auf E dann und nur dann gilt, wenn $\mathfrak{U}$ auch von links unitär ist. Diese „Besselsche Ungleichung" haben wir bei endlichen Matrizen nur bei den halbunitären Matrizen gefunden. Die halbunitären unendlichen Matrizen, zu welchen wir nun übergehen werden, sind allgemeiner als unendliche Matrizen, die von rechts unitär sind, aber nicht von links oder umgekehrt. Ist eine Matrix sowohl von rechts als auch von links unitär, so nennen wir sie schlechthin unitär. Hierfür ist notwendig und hinreichend

$$\mathsf{n}(\mathfrak{U}\mathfrak{U}^*) = \mathsf{n}(\mathfrak{U}^*\mathfrak{U}) = \mathsf{m}(\mathfrak{U}\mathfrak{U}^*) = \mathsf{m}(\mathfrak{U}^*\mathfrak{U}) = 1. \tag{255}$$

Die unitären sind diejenigen Matrizen, welche eine und nur eine Reziproke haben, die mit der begleitenden Matrix identisch ist, $\mathfrak{U}^{-1} = \mathfrak{U}^*$, wobei, wie gesagt, das Zeichen $^{-1}$ nur im Falle **1** benutzt werden soll, also im Falle **3**, mithin bei nur einseitig unitären Matrizen keinesfalls.

§ 69. Einzelmatrizen.

Es bezeiche γ_i, wie auf S. 57, eine Zahl, die entweder $=0$ oder $=1$ ist. Ist die hintere Norm $\mathfrak{U}\mathfrak{U}^*$ von $\mathfrak{U}$ von der Gestalt $\|\gamma_i \delta_{ik}\|$, so nennen wir $\mathfrak{U}$ halbunitär von rechts, und wir nennen $\mathfrak{U}$ halbunitär von links, wenn die vordere Norm $\mathfrak{U}^*\mathfrak{U}$ von dieser Gestalt ist. Wegen $\Phi(\|\gamma_i \delta_{ik}\|) \leqq \Phi(\mathfrak{E})$ sind die halbunitären Matrizen gewiß beschränkt und die Besselsche Ungleichheit (254), die für alle mindestens einseitig unitären Matrizen richtig ist, gilt wegen $\Phi(\|\gamma_i \delta_{ik}\|) \leqq \Phi(\|\delta_{ik}\|)$ erst recht für die halbunitären.

Eine beschränkte Hermitesche Matrix $\mathfrak{B}$ mit $\mathfrak{B}^2 = \mathfrak{B}$ nennen wir wieder eine Einzelmatrix. Offenbar ist $\|\gamma_i \delta_{ik}\|^2 = \|\gamma_i^2 \delta_{ik}^2\| = \|\gamma_i \delta_{ik}\|$. Es ist einleuchtend, daß die hintere Norm einer von rechts halbunitären sowie die vordere Norm einer von links halbunitären Matrix eine Einzelmatrix ist. Wie bei den endlichen Matrizen gibt es Einzelmatrizen, die nicht von der Gestalt $\|\gamma_i \delta_{ik}\|$ sind. Es gilt jedoch wieder der Satz, daß jede Einzelmatrix als die vordere Norm einer von rechts halbunitären oder als die hintere Norm einer von links halbunitären Matrix dargestellt werden kann. Dieser Hilbertsche Satz ist der einzige existentiale Satz algebraischer Natur in der Hauptachsentheorie der unendlichen Hermiteschen Matrizen; was noch hinzukommt, ist entweder eine Folgerung, wie z. B. die im nächsten Paragraphen zu besprechende Ergänzungsmöglichkeit zu einer unitären Matrix, oder eine Sache des in Kap. II entwickelten analytischen Apparats. — Was nun den Beweis des Hilbertschen Satzes anlangt, so brauchen wir darauf nicht näher einzugehen. In der Tat ist die in § 27 durchgeführte Jacobische Transformation derart, daß sie bei beschränkten unendlichen Matrizen wörtlich wiederholt werden kann. Wir haben ja die Zeilen der Matrix $\mathfrak{U}$, welche bei vorgegebenem $\mathfrak{B}$ die Eigenschaft $\mathfrak{B} = \mathfrak{U}^* \mathfrak{U}$, $\mathfrak{U}\mathfrak{U}^* = \|\gamma_i \delta_{ik}\|$ haben soll, *sukzessive* angegeben, und diese Konstruktion kann jetzt wiederholt werden. Man erkennt dies, wenn man die folgenden drei Tatsachen beachtet: 1. $\mathfrak{B}$ ist beschränkt, also gewiß eine Q-Matrix, so daß die Reihen $\sum_j |b_{ij}|^2$ konvergent sind; 2. $\mathfrak{B}$ ist Hermitesch, $= \mathfrak{B}^*$, und außerdem eine Einzelmatrix, $= \mathfrak{B}^2$, mithin $= \mathfrak{B}\mathfrak{B}^*$, also eine Norm und daher nichtnegativ definit; 3. die Abschnitte einer nichtnegativ definiten Matrix sind erst recht nichtnegativ definit. — Die Schursche Transformation ist in der auf S. 21 gegebenen Beweisanordnung hier nicht valent, und zwar eigentlich nicht darum, weil man zum Schluß unendlich viele Matrizen miteinander zu multiplizieren hätte (S. 22), sondern weil auf S. 21 der Fundamentalsatz der Algebra (ja überhaupt eine Säkular*determinante*) unter Einschaltung des Alternativsatzes benutzt wurde.

§ 70. Orthogonale Einzelmatrizen.

Bevor wir nun zu den bereits erwähnten Ergänzungsbetrachtungen übergehen, wollen wir noch einige von Hilbert ebenfalls benutzte elementare Bemerkungen über die Einzelmatrizen vorausschicken. Man sagt, die beschränkte Matrix $\mathfrak{C}$ sei zu einer anderen, $\mathfrak{D}$, orthogonal, wenn das Produkt $\mathfrak{C}\mathfrak{D} = \mathfrak{O} =$ Nullmatrix ist. Daraus folgt noch nicht, daß umgekehrt $\mathfrak{D}$ zu $\mathfrak{C}$ orthogonal ist (dies gilt nicht einmal bei den

endlichen Matrizen). Sind aber $\mathfrak{C}$ und $\mathfrak{D}$ Einzelmatrizen, so sind die beiden Gleichungen $\mathfrak{C}\mathfrak{D} = \mathfrak{O}$, $\mathfrak{D}\mathfrak{C} = \mathfrak{O}$ gleichwertig; m. a. W., orthogonale Einzelmatrizen sind immer vertauschbar. Denn bereits orthogonale Hermitesche Matrizen sind vertauschbar. Aus $\mathfrak{C}\mathfrak{D} = \mathfrak{O}$ folgt nämlich immer $(\mathfrak{C}\mathfrak{D})^* = \mathfrak{O}^*$, d. h. $\mathfrak{D}^*\mathfrak{C}^* = \mathfrak{O}$, und jetzt ist $\mathfrak{D}^* = \mathfrak{D}$, $\mathfrak{C}^* = \mathfrak{C}$. — Sind $\mathfrak{C}$ und $\mathfrak{D}$ Einzelmatrizen, so ist $\mathfrak{C} \pm \mathfrak{D}$ wegen

$$(\mathfrak{C} \pm \mathfrak{D})^2 = \mathfrak{C}^2 \pm \mathfrak{C}\mathfrak{D} \pm \mathfrak{D}\mathfrak{C} + \mathfrak{D}^2 = \mathfrak{C} + \mathfrak{D} \pm (\mathfrak{C}\mathfrak{D} + \mathfrak{D}\mathfrak{C})$$

dann und nur dann eine Einzelmatrix, wenn $\mathfrak{C}\mathfrak{D} + \mathfrak{D}\mathfrak{C} = \mathfrak{O}$ gilt, also gewiß dann, wenn $\mathfrak{C}$ und $\mathfrak{D}$ orthogonal sind. Ist $\mathfrak{C}$ eine Einzelmatrix, so ist es wegen $(\mathfrak{E} - \mathfrak{C})^2 = \mathfrak{E} - 2\,\mathfrak{C} + \mathfrak{C} = \mathfrak{E} - \mathfrak{C}$ auch $\mathfrak{E} - \mathfrak{C}$. Die Kopplungsform einer Einzelmatrix ist auf E stets $\leqq 1$ und $\geqq 0$. Denn jede Einzelmatrix ist eine Norm einer halbunitären Matrix, also ist sie gewiß nichtnegativ definit (was bereits erwähnt wurde) und auf E mit Rücksicht auf die Besselsche Ungleichung (254) nicht größer als 1. Ist $\mathfrak{C}$ eine Einzelmatrix, so sind die beiden („komplementären") Einzelmatrizen $\mathfrak{C}$, $\mathfrak{E} - \mathfrak{C}$ wegen $\mathfrak{E}\mathfrak{C} - \mathfrak{C}^2 = \mathfrak{C} - \mathfrak{C} = \mathfrak{O}$ gewiß orthogonal.

Ist $\mathfrak{U}$ eine unitäre Matrix, so bezeichnen wir $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ wieder als die Transformierte von $\mathfrak{A}$ durch $\mathfrak{U}$. Das Produkt zweier unitären Substitutionen ist wieder unitär und man hat $(\mathfrak{U}_1\mathfrak{U}_2)^{-1} = \mathfrak{U}_2^{-1}\mathfrak{U}_1^{-1} = \mathfrak{U}_2^*\mathfrak{U}_1^*$. Denn diese Matrix ist eine sowohl vordere als auch hintere Reziproke von $\mathfrak{U}_1\mathfrak{U}_2$. Ist $\mathfrak{U}$ unitär und $\mathfrak{C}$ eine Einzelmatrix, so ist auch die Transformierte $\mathfrak{U}\mathfrak{C}\mathfrak{U}^{-1}$ eine Einzelmatrix, wegen $(\mathfrak{U}\mathfrak{C}\mathfrak{U}^{-1})^2 = \mathfrak{U}\mathfrak{C}^2\mathfrak{U}^{-1} = \mathfrak{U}\mathfrak{C}\mathfrak{U}^{-1}$. Transformiert man kogredient zwei orthogonale Einzelmatrizen $\mathfrak{C}$, $\mathfrak{D}$ unitär, so ist das transformierte Paar wegen $(\mathfrak{U}\mathfrak{C}\mathfrak{U}^{-1})(\mathfrak{U}\mathfrak{D}\mathfrak{U}^{-1}) = \mathfrak{U}\mathfrak{C}\mathfrak{D}\mathfrak{U}^{-1} = \mathfrak{U}\mathfrak{O}\mathfrak{U}^{-1} = \mathfrak{O}$ ebenfalls orthogonal. Transformiert man daher Einzelmatrizen $\mathfrak{C}_1, \mathfrak{C}_2, \ldots$, die paarweise orthogonal sind, so daß also $\pm\,\mathfrak{C}_1 \pm \mathfrak{C}_2 \pm \cdots$ eine Einzelmatrix ist, mit einer unitären Matrix $\mathfrak{U}$, so ist auch die transformierte Summe $\pm\,\mathfrak{U}\mathfrak{C}_1\mathfrak{U}^{-1} \pm \mathfrak{U}\mathfrak{C}_2\mathfrak{U}^{-1} \pm \cdots$ eine Einzelmatrix. Es ist dabei stillschweigend berücksichtigt worden, daß der Hermitesche Charakter bei unitären Transformationen erhalten bleibt. Doch läßt sich dies ebenso wie bei den endlichen Matrizen (S. 30) verifizieren.

Sind $\mathfrak{C}_1, \mathfrak{C}_2, \ldots, \mathfrak{C}_m, \ldots$ unendlich viele, paarweise orthogonale Einzelmatrizen, so ist $0 \leqq \Phi\left(\sum_{m=1}^{n}\mathfrak{C}_m\right) = \sum_{m=1}^{n}\Phi(\mathfrak{C}_m) \leqq 1$, da die Summe von orthogonalen Einzelmatrizen eine Einzelmatrix ist; wegen $\Phi(\mathfrak{C}_m) \geqq 0$ ist daher die Reihe $\sum_{m=1}^{\infty}\Phi(\mathfrak{C}_m)$ in jedem Punkte von E konvergent und $\leqq 1$. Für uns genügt es zu wissen, daß die Summe unendlich

vieler paarweise orthogonaler Einzelmatrizen immer und eindeutig gebildet werden kann und eine beschränkte Hermitesche Matrix darstellt, die außerdem, wie man leicht beweist, eine Einzelmatrix ist. Es genügt dabei zu beachten, daß wegen $\Phi(\mathfrak{E}_m) \geqq 0$ die konvergente Reihe $\sum \Phi(\mathfrak{E}_m) \leqq 1$ absolut konvergent ist. Es wäre jedoch verfehlt, zu glauben, daß auch $\sum \mathsf{M}(\mathfrak{E}_m)$ zu konvergieren braucht.

§ 71. Die Trennung. Hauptsatz über Einzelmatrizen.

Es sei $\mathfrak{X}$ eine beliebige Matrix, $\mathfrak{Y}$ eine andere. Ist i_0 derart, daß in der i_0-ten Zeile oder in der i_0-ten Kolonne zum mindesten ein Element von $\mathfrak{X}$ nicht verschwindet, so mögen alle Elemente sowohl der i_0-ten Zeile als auch in der i_0-ten Kolonne von $\mathfrak{Y}$ verschwinden, und es soll entsprechendes gelten, auch wenn umgekehrt die nicht verschwindenden Elemente von $\mathfrak{Y}$ betrachtet werden. Wir sagen dann, die beiden Matrizen seien *getrennt*. Diese Bedingung ist offenbar hinreichend für $\mathfrak{X}\mathfrak{Y} = \mathfrak{Y}\mathfrak{X} = \mathfrak{O}$; notwendig ist sie durchaus nicht.

Z. B. liegen die wie „irreduzible Quadrate“ (vgl. S. 146) gelegenen Matrizenbereiche getrennt; umgekehrt kann jede getrennt dargestellte endliche Matrix mittels einer Permutationsmatrix in eine Gesamtheit getrennter Quadrate übergeführt werden. Bei unendlichen Matrizen ist der Sachverhalt insofern, aber auch *nur* insofern komplizierter, als bereits das erste Quadrat unendlich sein kann. So muß man für die weiteren „Quadrate“ inmitten des ersten Quadrats „Platz machen“.

Es sei $\mathfrak{E}_1$ eine Einzelmatrix, mithin $\mathfrak{E}_0 = \mathfrak{E} - \mathfrak{E}_1$ eine zu $\mathfrak{E}_1$ orthogonale Einzelmatrix. Es gibt also von rechts halbunitäre Matrizen $\mathfrak{V}_0, \mathfrak{V}_1$ derart, daß $\mathfrak{E}_0 = \mathfrak{V}_0^* \mathfrak{V}_0$, $\mathfrak{E}_1 = \mathfrak{V}_1^* \mathfrak{V}_1$ gilt. Um beide Matrizen voneinander zu „trennen“, schaffen wir den hierfür erforderlichen Nullen zunächst dadurch „Platz“, daß wir von $\mathfrak{V}_0, \mathfrak{V}_1$ zu den offenbar getrennten (also orthogonalen) Matrizen

$$\tilde{\mathfrak{V}}_0\colon\ \tilde{v}^{(0)}_{2i\,2k} = v^{(0)}_{ik},\quad \tilde{v}^{(0)}_{2i+1\,2k+1} = \tilde{v}^{(0)}_{2i\,2k+1} = \tilde{v}^{(0)}_{2i+1\,2k} = 0;\quad i,k = 1, 2, \ldots,$$

$$\tilde{\mathfrak{V}}_1\colon\ \tilde{v}^{(1)}_{2i+1\,2k+1} = v^{(1)}_{ik},\quad \tilde{v}^{(1)}_{2i\,2k} = \tilde{v}^{(1)}_{2i\,2k+1} = \tilde{v}^{(1)}_{2i+1\,2k} = 0;\quad i,k = 1, 2, \ldots$$

übergehen. Durch Zusammenfassung ergibt sich daraus (vgl. die Ergänzungsbetrachtungen in § 27) eine unitäre Matrix, d. h. ihre begleitende Matrix ist ihre einzige Reziproke (während $\mathfrak{V}_0, \mathfrak{V}_1$ zwar von rechts halbunitär, jedoch vielleicht nicht einmal von rechts unitär waren). Bezeichnet $\mathfrak{U}$ diese unitäre Matrix, so sind die Matrizen $\mathfrak{U}\mathfrak{E}_0\mathfrak{U}^{-1}$, $\mathfrak{U}\mathfrak{E}_1\mathfrak{U}^{-1}$, wie wir wissen, gewiß orthogonale Einzelmatrizen, und darüber hinaus sind sie jetzt, wie man sich leicht überzeugt, von-

einander getrennt. — Man beachte dabei, daß die Einzelmatrizen nichtnegativ definit sind und daher außer bei der Nullmatrix zum mindesten ein Diagonalelement ($\geqq 0$) gewiß $\neq 0$ ist.

Man überblickt die Verhältnisse am einfachsten, wenn man zu den Kopplungsformen übergeht: zwei Matrizen $\mathfrak{A}$, $\mathfrak{B}$ sind dann und nur dann getrennt, wenn

$$(257)\quad \sum_{i=1}^{\infty}\sum_{k=1}^{\infty}(a_{ik}+b_{ik})\,x_i\bar{x}_k \equiv \sum_{i=1}^{\infty}\sum_{k=1}^{\infty}a_{ik}\,x_i'\bar{x}_k' + \sum_{i=1}^{\infty}\sum_{k=1}^{\infty}b_{ik}\,x_i''\bar{x}_k''$$

gesetzt werden kann, wobei jedes x_i mit einem und nur einem x_i' *oder* x_i'' identisch ist und umgekehrt, so daß kein x_i' mit einem x_i'' identisch ist und umgekehrt.

Wir haben vorher nur eine Einzelmatrix $\mathfrak{E}_1$ betrachtet und dann das orthogonale Komplement $\mathfrak{E}_0 = \mathfrak{E} - \mathfrak{E}_1$. Wählt man beim „Platzmachen" die „Plätze" auf eine leicht ersichtliche Weise „breiter", so ergibt sich ebenso folgendes. Es seien $\mathfrak{E}_1, \mathfrak{E}_2, \ldots, \mathfrak{E}_m, \ldots$ paarweise orthogonale Einzelmatrizen und $\mathfrak{E}_0 = \mathfrak{E} - \sum \mathfrak{E}_m$ das Komplement ihrer Summe, das gewiß zu jedem $\mathfrak{E}_m$ orthogonal ist. Denn es ist zunächst $\mathfrak{E}_0 = \mathfrak{E} - \mathfrak{E}_{m_0} - \sum' \mathfrak{E}_m$, woselbst $\mathfrak{E}_{m_0}$ ein bestimmtes $\mathfrak{E}_m$ und $\sum' \mathfrak{E}_m$ die Summe aller übrigen bezeichnet. Nun ist $\sum'$ gewiß zu $\mathfrak{E}_{m_0}$ orthogonal und $\mathfrak{E}_{m_0}$ ist zu $\mathfrak{E} - \mathfrak{E}_{m_0}$ ebenfalls orthogonal. Offenbar kann man daher die vorige Betrachtung auch dann wiederholen, wenn die Anzahl der Summanden unendlich ist. Denn bei dem „Platzmachen" ergeben sich, da man nicht alle Nullenrechtecke von der gleichen Größe zu wählen braucht, keine Schwierigkeiten, und auch Konvergenzschwierigkeiten liegen bei orthogonalen Einzelmatrizen nach § 70 nie vor, also erst recht dann nicht, wenn sie getrennt sind oder getrennt werden. Es ergibt sich so der folgende

Hauptsatz über Einzelmatrizen. Es sei $\mathfrak{E}_1, \mathfrak{E}_2, \ldots$ eine endliche oder unendliche Folge von paarweise orthogonalen Einzelmatrizen, $\mathfrak{E}_0 = \mathfrak{E} - \sum \mathfrak{E}_m$ das zu allen $\mathfrak{E}_m$ gewiß orthogonale Komplement der Summe, so daß es also von rechts halbunitäre Matrizen $\mathfrak{B}_0, \mathfrak{B}_1, \ldots$ gibt, für welche $\mathfrak{E}_0 = \mathfrak{B}_0^* \mathfrak{B}_0$, $\mathfrak{E}_1 = \mathfrak{B}_1^* \mathfrak{B}_1, \ldots$ gilt. Es existiert dann eine unitäre Matrix $\mathfrak{U}$ derart, daß die Matrizen $\mathfrak{U}\,\mathfrak{E}_0\,\mathfrak{U}^{-1}$, $\mathfrak{U}\,\mathfrak{E}_1\,\mathfrak{U}^{-1}, \ldots$ paarweise getrennt sind. Da wegen des Getrenntseins für eine gliedweise Transformation von $\sum \mathfrak{E}_m$ keinerlei Schwierigkeiten vorliegen, so gilt also

$$(256)\qquad \mathfrak{E} = \mathfrak{U}\,\mathfrak{E}_0\,\mathfrak{U}^{-1} + \mathfrak{U}\,\mathfrak{E}_1\,\mathfrak{U}^{-1} + \cdots,$$

wobei die Summanden getrennte Diagonalmatrizen sind, deren nichtverschwindende Elemente $= 1$ ausfallen.

Es ist nämlich, wenn $\mathfrak{W}_j = \mathfrak{U}\,\mathfrak{B}_j^*\,\mathfrak{U}^{-1}$ gesetzt wird, wegen $\mathfrak{C}_j = \mathfrak{B}_j^*\,\mathfrak{B}_j$ der transformierte Summand $\mathfrak{U}\,\mathfrak{C}_j\,\mathfrak{U}^{-1} = \mathfrak{W}_j\,\mathfrak{W}_j^*$. Nun ist aber $\mathfrak{B}_j\,\mathfrak{B}_j^*$ nach Voraussetzung von der Gestalt $\|\gamma_i^{(j)}\,\delta_{ik}\|$, wobei $\gamma_i^{(j)}$ entweder $= 0$ oder $= 1$ ist, und die Matrix $\mathfrak{U}$ ist eben deshalb vermittelst der Trennung, (S. 158) erklärt worden damit die $\mathfrak{W}_j\,\mathfrak{W}_j^*$ ebenfalls getrennt liegen.

§ 72. Spektrale Einzelmatrizen. Die drei Bestandteile des Spektrums.

Es sei $\|\sigma_{ik}(\mu)\|$ eine beschränkte Hermitesche Matrix der für $-\infty < \mu < +\infty$ erklärten, der Bedingung $\sigma_{ik}(-\infty) = 0$ genügenden, von rechts stetigen Funktionen, die von gleichmäßig beschränkter Schwankung, also auch gleichmäßig beschränkt sind. Es sei ferner die Kopplungsform $\Phi(\|\sigma_{ik}(\mu)\|;\,\mathfrak{x}) = \sum\sum \sigma_{ik}(\mu)\,x_i\,\bar{x}_k$ in jedem festen Punkte von E eine nicht abnehmende Funktion von μ. Insbesondere ist also $\mathsf{M}(\|\sigma_{ik}(\mu)\|)$ nicht abnehmend. Wir nehmen an, daß $\mathsf{M}(\|\sigma_{ik}(+\infty)\|)$ endlich, also die Funktionenmatrix gleichmäßig beschränkt ist. Es sei endlich

$$(258)\qquad \|\sigma_{ik}(\mu)\|\,\|\sigma_{ik}(\nu)\| = \begin{cases} \|0\| = \mathfrak{O} & \text{für}\quad \mu \neq \nu, \\ \|\sigma_{ik}(\mu)\| = \|\sigma_{ik}(\nu)\| & \text{für}\quad \mu = \nu, \end{cases}$$

also $\|\sigma_{ik}(\mu)\|$ eine zu $\|\sigma_{ik}(\nu)\|$, $\mu \neq \nu$ orthogonale Einzelmatrix. Sind alle diese Bedingungen erfüllt, so sagen wir, die Funktionenmatrix $\|\sigma_{ik}(\mu)\|$ sei eine spektrale Einzelmatrix. Aus (258) folgt leicht unter Benutzung der auf S. 52 eingeführten Schreibweise

$$(259)\qquad \|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\|\,\|\underset{\mathrm{b}}{\Delta}\sigma_{ik}\| = \|\underset{\mathrm{ab}}{\Delta}\sigma_{ik}\|,$$

sofern zunächst a und b Intervalle bezeichnen, deren Endpunkte Stetigkeitsstellen aller $\sigma_{ik}(\mu)$ sind. Indem man jedoch a oder b oder beide auf einen Punkt zusammenschrumpfen läßt, folgt aus (258) unter Benutzung des Majorantenkriteriums (S. 134) sowie etwa des Umstandes, daß $\mathsf{M}(\|\sigma_{ik}(\mu)\|)$ eine nicht abnehmende Funktion ist, daß (259) auch dann gilt, wenn a oder b oder beide beliebige Punkte bezeichnen, und daraus folgt endlich, daß (259) auch dann gilt, wenn Intervalle zugelassen werden, deren Endpunkte Sprungstellen sind. Mithin ist $\|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\|$ ohne Einschränkung eine zu $\|\underset{\mathrm{b}}{\Delta}\sigma_{ik}\|$, $\mathrm{a} \neq \mathrm{b}$, orthogonale Einzelmatrix.

Den uninteressanten Fall $\mathsf{M}(\|\sigma_{ik}(+\infty)\|) = 0$, wo jedes $\sigma_{ik}(\mu)$ für alle μ verschwindet, können wir ausschließen. Wir sagen, der Punkt μ_0 gehöre dem Spektrum von $\|\sigma_{ik}\|$ an, wenn es kein $\varepsilon > 0$ gibt derart, daß $\|\sigma_{ik}(\mu)\|$ auf dem Intervall $-\varepsilon + \mu_0 < \mu < \mu_0 + \varepsilon$ konstant, d. h. $\equiv \|\sigma_{ik}(\mu_0)\|$ ist. Nach der soeben getroffenen Verabredung enthält das Spektrum mindestens einen Punkt, es kann aber auch ein endliches oder unendliches Intervall, allgemeiner eine beliebige abgeschlossene Punktmenge der μ-Gerade sein, während es andererseits offenbar immer abgeschlossen ist. Ist der Punkt μ_0 derart, daß mindestens ein $\underset{\mu_0}{\Delta}\sigma_{ik}$ von Null verschieden ausfällt, so sagen wir, μ_0 gehöre dem Punktspektrum an, das auch leer sein kann. Für alle Fälle gibt es, da jedes σ_{ik} höchstens abzählbar viele Sprungstellen haben kann (S. 75 u. S. 80), im Punktspektrum höchstens abzählbar viele Punkte $\lambda_1, \lambda_2, \ldots$, die jedoch auch überall dicht liegen können. Führen wir die reinen Sprungfunktionen

$$\hat{\sigma}_{ik}(\mu) = \sum_{\lambda_j \leqq \mu} \underset{\lambda_j}{\Delta}\sigma_{ik}$$

ein, so sind die Funktionen

$$\tilde{\sigma}_{ik}(\mu) = \sigma_{ik}(\mu) - \hat{\sigma}_{ik}(\mu)$$

durchweg stetig (S. 76). Die Gesamtheit der Punkte $\tilde{\mu}_0$, in welchen es kein $\varepsilon > 0$ mit $\|\tilde{\sigma}_{ik}(\mu)\| \equiv \|\tilde{\sigma}_{ik}(\tilde{\mu}_0)\|$ für $-\varepsilon + \tilde{\mu}_0 < \mu < \tilde{\mu}_0 + \varepsilon$ gibt, nennen wir Streckenspektrum von $\|\sigma_{ik}\|$. Diese Punktmenge ist, wenn sie überhaupt Punkte enthält, wieder abgeschlossen und darüber hinaus perfekt, d. h. mit der Menge ihrer Häufungsstellen identisch. Denn die $\tilde{\sigma}_{ik}(\mu)$ sind stetig, so daß jeder Punkt $\tilde{\mu}_0$ Häufungspunkt von Punkten $\tilde{\mu}_0$ sein muß. Punktspektrum und Streckenspektrum gehören gewiß dem Spektrum an, können eventuell übereinandergreifen, doch erschöpfen sie das Spektrum eventuell noch nicht. Da nämlich das Punktspektrum im allgemeinen nicht abgeschlossen ist, so kann es Häufungsstellen haben, die weder dem Punktspektrum noch dem Streckenspektrum angehören. Die Gesamtheit der (eventuellen) Häufungsstellen des Punktspektrums, die auch dem Punkt-, aber auch dem Streckenspektrum angehören können, nennen wir Häufungsspektrum.

§ 73. Fortsetzung. Anwendung des Hauptsatzes über Einzelmatrizen.

Wegen (259) ist $\|\underset{\lambda_1}{\Delta}\sigma_{ik}\|, \|\underset{\lambda_2}{\Delta}\sigma_{ik}\|, \ldots$ eine Folge von paarweise orthogonalen Einzelmatrizen, also ist auch $\|\hat{\sigma}_{ik}(\mu)\| = \|\sum_{\lambda_l \leqq \mu} \underset{\lambda_l}{\Delta}\sigma_{ik}\|$ eine

spektrale Einzelmatrix, die zu allen Einzelmatrizen $\|\underset{\lambda_j}{\Delta}\sigma_{ik}\|$ orthogonal ist, sobald $\mu \neq \lambda_j$ ausfällt. Aus (258) kann ferner, analog zu (259), unmittelbar entnommen werden, daß auch die Matrix $\|\tilde{\sigma}_{ik}(\mu)\| = \|\sigma_{ik}(\mu) - \hat{\sigma}_{ik}(\mu)\|$ eine spektrale Einzelmatrix ist, die zu $\|\underset{\lambda_j}{\Delta}\sigma_{ik}\|$ und daher zu $\|\hat{\sigma}_{ik}(\nu)\|$, $\nu \neq \mu$ orthogonal ist. Es ist also

$$(260) \qquad \|\sigma_{ik}(\mu)\| = \|\tilde{\sigma}_{ik}(\mu)\| + \sum_{\lambda_j \leqq \mu} \|\underset{\lambda_j}{\Delta}\sigma_{ik}\|,$$

wobei die rechterhand stehenden Einzelmatrizen paarweise orthogonal sind, wenn μ nicht dem Punktspektrum angehört. Da aber die Gleichung

$$(261) \qquad \|\tilde{\sigma}_{ik}(\mu)\|\ \|\underset{\lambda_j}{\Delta}\sigma_{ik}\| = \|0\|$$

für $\mu \neq \lambda_j$ gewiß richtig ist, so muß sie auch für $\mu = \lambda_j$ gelten. Man erkennt dies wieder nach Vornahme des Grenzüberganges $\mu \rightarrow \lambda_j$ und unter Beachtung des Majorantenkriteriums. Folglich sind die in (260) rechterhand stehenden Matrizen, und übrigens offenbar ebenso auch die beiden Matrizen $\|\tilde{\sigma}_{ik}(\mu)\|$, $\|\hat{\sigma}_{ik}(\mu)\|$, auch dann orthogonal, wenn μ dem Punktspektrum angehört. Es gibt also nach dem Hauptsatz über Einzelmatrizen eine unitäre Matrix $\mathfrak{U}$ derart, daß

$$(262) \qquad \mathfrak{U}\|\sigma_{ik}(\mu)\|\mathfrak{U}^{-1} = \mathfrak{U}\|\tilde{\sigma}_{ik}(\mu)\|\mathfrak{U}^{-1} + \sum_{\lambda_j \leqq \mu} \mathfrak{U}\|\underset{\lambda_j}{\Delta}\sigma_{ik}\|\mathfrak{U}^{-1}$$

gilt, wobei die rechterhand stehenden Einzelmatrizen

$$(263) \qquad \mathfrak{U}\|\tilde{\sigma}_{ik}(\mu)\|\mathfrak{U}^{-1} = \mathfrak{W}(\mu)\,\mathfrak{W}^*(\mu), \qquad \mathfrak{U}\|\underset{\lambda_j}{\Delta}\sigma_{ik}\|\mathfrak{U}^{-1} = \mathfrak{W}_j\,\mathfrak{W}_j^*$$

diesmal nicht nur orthogonal, sondern sogar getrennt sind; auch die Matrizen $\mathfrak{W}(\mu)$, $\mathfrak{W}_1$, $\mathfrak{W}_2$, ... sind dabei getrennt und außerdem von rechts halbunitär, d. h. es ist $\mathfrak{W}_j\,\mathfrak{W}_j^* = \|\gamma_i^{(j)}\delta_{ik}\|$, $\mathfrak{W}(\mu)\,\mathfrak{W}^*(\mu) = \|\gamma_i(\mu)\,\delta_{ik}\|$, wobei die γ entweder $=0$ oder $=1$ sind. Es ist also

$$\mathfrak{U}\|\sigma_{ik}(\mu)\|\mathfrak{U}^{-1} = \|\gamma_i(\mu)\,\delta_{ik}\| + \sum_{\lambda_j \leqq \mu}\|\gamma_i^{(j)}\delta_{ik}\|,$$

wobei die Matrizen rechterhand paarweise getrennt sind. Dabei ist zu beachten, daß die Matrix $\mathfrak{U}$ noch von μ abhängt. Doch kann die trennende „Transformation auf die Hauptachsen" bei dem punktspektralen Bestandteil offenbar auch unabhängig von μ gewählt werden, d. h. derart, daß die $\gamma_i^{(j)}$ unabhängig von μ werden (man erkennt dies, wenn man in (260) einfach über alle j summiert). Da andererseits die Matrix $\|\tilde{\sigma}_{ik}(\mu)\|$, welche in $\|\gamma_i(\mu)\,\delta_{ik}\|$ übergeführt wird, von μ wohl abhängig ist, so haben wir von vornherein $\mathfrak{U}$ für $\mathfrak{U}(\mu)$ geschrieben. — Es ist dabei ein

springender Punkt [der den Umstand, daß nicht jede Hermitesche Matrix auf die Diagonalform gebracht werden kann (vgl. S. 201), überhaupt verständlich macht], daß $\gamma_i(\mu)$ in der Diagonalmatrix $\|\gamma_i(\mu)\,\delta_{ik}\| = \mathfrak{U}\,\|\tilde{\sigma}_{ik}(\mu)\|\,\mathfrak{U}^{-1}$ zwar nur der beiden Werte 0, 1 fähig ist, daß es aber von μ abhängt, ob $\gamma_i(\mu) = 0$ oder aber $= 1$ ausfällt und daß überhaupt $\mathfrak{U}$ von μ abhängig ist. Es handelt sich also um eine (im Streckenspektrum) *lokale*, „infinitesimale" Normalgestalt [und mehr kann, sobald es ein Streckenspektrum gibt, überhaupt nicht erreicht werden]. — Wir können also nur folgendes behaupten. Zu jedem μ_0 gibt es eine unitäre Matrix $\mathfrak{U}$ derart, daß, wenn $\|\tilde{\sigma}^*_{ik}(\mu)\| = \mathfrak{U}\,\|\tilde{\sigma}_{ik}(\mu)\|\,\mathfrak{U}^{-1}$ gesetzt wird,

$$(264)\qquad \mathfrak{U}\,\|\sigma_{ik}(\mu)\|\,\mathfrak{U}^{-1} = \|\tilde{\sigma}^*_{ik}(\mu)\| + \sum_{\lambda_j \leqq \mu} \|\gamma_i^{(j)}\,\delta_{ik}\|$$

gilt, wobei die $\gamma_i^{(j)}$ unabhängig von μ ausfallen und die Matrizen rechterhand für $\mu = \mu_0$ gewiß paarweise getrennt sind. Insbesondere sind also die zu verschiedenen j gehörigen $\|\gamma_i^{(j)}\,\delta_{ik}\|$ voneinander gewiß getrennt.

Wir behaupten nun, daß die beiden Matrizen $\|\tilde{\sigma}^*_{ik}(\mu)\|$, $\|\gamma_i^{(j)}\,\delta_{ik}\|$ nicht nur für $\mu = \mu_0$, sondern für alle $\mu < \mu_0$ getrennt sein müssen. In unseren früheren Bezeichnungen ist nämlich $\tilde{\sigma}^*_{ik}(\mu_0) = \gamma_i(\mu_0)\,\delta_{ik}$. Da $\Phi(\|\tilde{\sigma}^*_{ik}(\mu)\|)$ mit μ nicht abnimmt und $\geqq 0$ ist, so folgt, indem man darin $x_i = 1$ und die übrigen $x = 0$ setzt, $0 \leqq \tilde{\sigma}^*_{ii}(\mu) \leqq \tilde{\sigma}^*_{ii}(\mu_0) = \gamma_0(\mu_0)$, wenn $\mu < \mu_0$. Ist also $\gamma_i(\mu_0) = 0$, so muß $\tilde{\sigma}^*_{ii}(\mu) = 0$, also, mit Rücksicht darauf, daß die Einzelformen nichtnegativ definit sind, auch $\tilde{\sigma}^*_{ik}(\mu) = \tilde{\sigma}^*_{ki}(\mu) = 0$; $k = 1, 2, \ldots$ gelten [vgl. nämlich (60)], woraus die Behauptung folgt. Gibt es also eine Zahl M derart, daß für $\mu > M$ alle $\tilde{\sigma}^*_{ik}(\mu)$ konstant, $= \tilde{\sigma}^*_{ik}(M)$ sind, so gibt es eine von μ unabhängige unitäre Matrix $\mathfrak{U}$ derart, daß in (264) die Matrizen rechterhand für alle μ getrennt sind.

§ 74. Fortsetzung. Unitäre Kovarianz. Invarianz der Spektra.

Ist $\|\sigma_{ik}(\mu)\|$ eine spektrale Einzelmatrix, so ist es auch $\mathfrak{U}\,\|\sigma_{ik}(\mu)\|\,\mathfrak{U}^{-1}$, wenn $\mathfrak{U}$ unitär ist und von μ nicht abhängt. Aus (258) folgt nämlich nach den Faltungssätzen, die angewandt werden können, weil jede unitäre Matrix und jede Einzelmatrix beschränkt ist, daß

$$(265)\qquad (\mathfrak{U}\,\|\sigma_{ik}(\mu)\|\,\mathfrak{U}^{-1})(\mathfrak{U}\,\|\sigma_{ik}(\nu)\|\,\mathfrak{U}^{-1}) = \begin{cases} \|0\| = \mathfrak{O} & \text{für } \mu \neq \nu \\ \mathfrak{U}\,\|\sigma_{ik}(\mu)\|\,\mathfrak{U}^{-1} & \text{für } \mu = \nu \end{cases}$$

gilt, und die unitäre Invarianz der weiteren, von den Einzelmatrizen auf

S. 160 geforderten Eigenschaften ist ebenfalls einleuchtend. Wir setzen

$$\sigma_{ik}^{*}(\mu) = \mathfrak{U} \, \| \sigma_{ik}(\mu) \| \, \mathfrak{U}^{-1} \tag{266}$$

und bezeichnen entsprechend mit $\| \hat{\sigma}_{ik}^{*}(\mu) \|$ den rein sprunghaften, mit $\| \tilde{\sigma}_{ik}^{*}(\mu) \|$ den stetigen Bestandteil von $\| \sigma_{ik}^{*}(\mu) \|$. Wir wollen zeigen, daß $\sigma_{ik}^{*}(\mu)$ von rechts stetig ist. Die Funktionen $\sigma_{ik}(\mu)$ sind nämlich nach Voraussetzung von rechts stetig, also ist es auch die Funktion $\xi_{jk}(\mu) = \sum_{l} \sigma_{jl}(\mu)\, \bar{u}_{kl}$. Denn einerseits ist die Matrix $\| u_{ik} \|$ unitär, mithin beschränkt, folglich eine Q-Matrix, also ist $\sum_{l} x_l \, \bar{u}_{kl}$ bei jedem k eine beschränkte Linearform und daher gleichmäßig konvergent für alle x_l mit $\sum_{l} | x_l |^2 < \text{Konst.}$; siehe S. 127. Andererseits ist $\| \sigma_{jl}(\mu) \|$ eine Einzelmatrix, also ist ihr P gewiß $\leqq 1$, mithin ist $\sum_{l} | \sigma_{jl}(\mu) |^2 \leqq 1$ bei jedem j. Da aber eine gleichmäßig konvergente Reihe von rechtsstetigen Funktionen bekanntlich von rechts stetig ist, so sind die $\xi_{jk}(\mu)$ von rechts stetig. Nach der Besselschen Ungleichheit ist ferner $\sum_{j} | \xi_{jk}(\mu) |^2 \leqq 1$, so daß eine Wiederholung der vorigen Schlußweise ergibt, daß $\sum_{j} u_{ij} \, \xi_{jk}(\mu)$, d. h. $\sigma_{ik}^{*}(\mu)$, wie behauptet, von rechts stetig ist. Ebenso erkennt man, daß

$$\lim_{\varepsilon \to 0} \sigma_{ik}^{*}(\mu \pm \varepsilon) = \sum_{j} \sum_{l} u_{ij} \lim_{\varepsilon \to 0} \sigma_{ik}(\mu \pm \varepsilon)\, \bar{u}_{kl}, \tag{267}$$

d. h. $\| \sigma_{ik}^{*}(\mu \pm 0) \| = \mathfrak{U} \, \| \sigma_{ik}(\mu \pm 0) \| \, \mathfrak{U}^{-1}$, mithin $\| \underset{a}{\Delta} \sigma_{ik}^{*} \| = \mathfrak{U} \, \| \underset{a}{\Delta} \sigma_{ik} \| \, \mathfrak{U}^{-1}$ gilt. — Das Spektrum, d. h. die Menge der Nichtkonstanzstellen von $\| \sigma_{ik}(\mu) \|$, ist selbstverständlich unitär invariant (denn $\mathfrak{U}^{-1}$ ist auch eine unitäre Matrix). Ist μ_0 eine Stetigkeitsstelle aller σ_{ik}, so ist μ_0 nach dem Obigen eine Stetigkeitsstelle aller σ_{ik}^{*}. Ist also μ_0 Sprungstelle von mindestens einem σ_{ik}, so gilt dies auch von mindestens einem σ_{ik}^{*}. Denn sonst würde eine Stetigkeitsstelle aller σ_{ik}^{*} mittels der unitären Substitution $\mathfrak{U}^{-1}$ in eine Sprungstelle von mindestens einem σ_{ik} übergehen. Es sind also auch das Streckenspektrum und das Punktspektrum und daher endlich auch das Häufungsspektrum unitär invariant, und dies ergibt offenbar

$$\begin{gathered} \| \tilde{\sigma}_{ik}^{*}(\mu) \| = \mathfrak{U} \, \| \tilde{\sigma}_{ik}(\mu) \| \, \mathfrak{U}^{-1}, \qquad \| \hat{\sigma}_{ik}^{*}(\mu) \| = \mathfrak{U} \, \| \hat{\sigma}_{ik}(\mu) \| \, \mathfrak{U}^{-1}, \\ \| \underset{\lambda_j}{\Delta} \sigma_{ik}^{*} \| = \mathfrak{U} \, \| \underset{\lambda_j}{\Delta} \sigma_{ik} \| \, \mathfrak{U}^{-1}. \end{gathered} \tag{268}$$

Unsere Ergebnisse können wir also mit Rücksicht auf § 73 wie folgt aussprechen:

Zu jeder spektralen Einzelmatrix, die für hinreichend große μ konstant ist, gibt es eine von μ unabhängige unitäre Substitution $\mathfrak{U}$ derart, daß die transformierte spektrale Einzelmatrix

$$\| \sigma_{ik}^*(\mu) \| = \mathfrak{U} \, \| \sigma_{ik}(\mu) \| \, \mathfrak{U}^{-1}$$

in

$$\| \sigma_{ik}^* \| = \| \tilde{\sigma}_{ik}^* \| + \| \hat{\sigma}_{ik}^* \|,$$

oder sogar in

$$\| \sigma_{ik}^*(\mu) \| = \| \tilde{\sigma}_{ik}^*(\mu) \| + \sum_{\lambda_j \leqq \mu} \| \underset{\lambda_j}{\varDelta} \sigma_{ik}^* \| \tag{269}$$

zerfällt werden kann; dabei sind die Matrizen rechterhand paarweise getrennte Matrizen derart, daß $\underset{\lambda_j}{\varDelta} \sigma_{ik}^* = 0$ ist für $i \neq k$, während $\underset{\lambda_j}{\varDelta} \sigma_{ii}^*$ entweder $= 0$ oder $= 1$ ist, je nachdem wie i und j beschaffen sind. Die λ_j sind Stellen des Punktspektrums von $\| \sigma_{ik}^* \|$, d. h. von $\| \sigma_{ik} \|$, und $\| \tilde{\sigma}_{ik}^* \|$ ist der stetige Bestandteil von $\| \sigma_{ik}^* \|$ oder die Transformierte des stetigen Bestandteiles von $\| \sigma_{ik} \|$; die Sprungmatrix $\| \underset{\lambda_j}{\varDelta} \sigma_{ik}^* \|$ ist die Transformierte der Sprungmatrix $\| \underset{\lambda_j}{\varDelta} \sigma_{ik} \|$ von $\| \sigma_{ik} \|$.

§ 75. Fortsetzung. Die Vorteile des Getrenntseins.

Um der Schwierigkeit, daß $\tilde{\sigma}_{ik}^*(\mu)$ noch von μ abhängt, gerecht zu werden, werden wir jetzt $\| \tilde{\sigma}_{ik}^*(\mu) \|$ nach Hellinger mittels weiterer Jacobischen Transformationen zerlegen. Wir unterwerfen also $\| \sigma_{ik}^*(\mu) \| = \mathfrak{U} \, \| \sigma_{ik}(\mu) \| \, \mathfrak{U}^{-1}$ geeigneten weiteren unitären Transformationen $\mathfrak{U}_1$, $\mathfrak{U}_2, \ldots$. Beachten wir, daß das Produkt unitärer Transformationen wieder unitär ist, so finden wir, daß $\| \sigma_{ik}(\mu) \|$ alles in allem nur einer unitären Transformation unterworfen wird. Wir werden dabei die bereits gewonnenen Normalformen von Bestandteilen durch die neuen unitären Transformationen keinesfalls zerstören. Wir werden z. B. zunächst $\| \tilde{\sigma}_{ik}^*(\mu) \|$ einer unitären Substitution $\mathfrak{U}_1$ unterwerfen, welche $\| \tilde{\sigma}_{ik}^*(\mu) \|$ in zwei getrennte Summanden zerlegt. Nun ist aber $\| \tilde{\sigma}_{ik}^*(\mu) \|$ von den $\| \underset{\lambda_j}{\varDelta} \sigma_{ik}^* \|$ bereits getrennt, so daß wir bei $\| \tilde{\sigma}_{ik}^*(\mu) \|$ alles Erreichbare auch durch eine unitäre Substitution ${}^0\mathfrak{U}_1$ erreichen können, in der das i-te Element der k-ten Zeile $= \delta_{ik}$ ist, wenn die i-te Zeile oder Kolonne oder die k-te Zeile oder Kolonne mindestens ein von Null verschiedenes $\underset{\lambda_j}{\varDelta} \sigma^*$ enthält; ${}^0\mathfrak{U}_1$ entsteht dabei aus $\mathfrak{U}_1$ einfach durch „Platzmachen“, wobei freilich in den Platz verschaffenden Quadraten die diagonalen Stellen (i, i) nicht mit 0, sondern mit 1 besetzt werden. Wir können uns daher Betrachtungen über Platzschaffen ein für allemal

ersparen. Wir können ja $\|\tilde{\sigma}_{ik}^*\|$ durch eine beliebige unitäre Substitution $\mathfrak{U}_1$ transformieren, denn es ist klar, daß wir sodann zu einer unitären Matrix ${}^0\mathfrak{U}_1$ übergehen können, welche für $\|\tilde{\sigma}_{ik}^*\|$ dasselbe wie $\mathfrak{U}_1$ leistet und dabei die bereits erreichte Normalform der Sprungmatrizen nicht zerstört. Wir brauchen also $\|\sigma_{ik}^*\|$ nicht mehr zu betrachten, sondern nur noch $\|\tilde{\sigma}_{ik}^*\|$. Wir können also durch $\mathfrak{U}_1$ von $\|\tilde{\sigma}_{ik}^*\|$ eine Matrix abspalten, so daß in der Zerlegung $\mathfrak{U}_1 \|\tilde{\sigma}_{ik}^*\| \mathfrak{U}_1^{-1} = \mathfrak{U}_1 \mathfrak{A} \mathfrak{U}_1^{-1} + \mathfrak{U}_1 \mathfrak{B} \mathfrak{U}_1^{-1}$ die beiden Summanden getrennt sind und $\mathfrak{U}_1 \mathfrak{A} \mathfrak{U}_1^{-1}$ eine Normalform von gewünschter Beschaffenheit hat. Wegen des Getrenntseins brauchen wir uns also um $\mathfrak{U}_1 \mathfrak{A} \mathfrak{U}_1^{-1}$ wieder nicht weiter zu kümmern und können aus $\mathfrak{U}_1 \mathfrak{B} \mathfrak{U}_1^{-1}$ durch eine beliebige unitäre Transformation $\mathfrak{U}_2$ einen weiteren Bestandteil von einer gewünschten Normalgestalt abspalten, ohne Gefahr zu laufen, daß dies eventuell die Preisgabe der Normalform der bereits abgespalteten Terme bedingt, usf. — Wir werden es sogar mit unendlich vielen Schritten zu tun haben, also zum Schluß unendlich viele unitäre Matrizen multiplizieren müssen, was jedoch, dank des sukzessiven Getrenntseins, offenbar keinerlei Schwierigkeit macht; das unendliche Produkt ist wieder eine unitäre Matrix und man kann mit ihm so rechnen, also ob es nur endlichviele Faktoren hätte. — Wir haben diese selbstverständlichen Bemerkungen vorausgeschickt, um spätere Betrachtungen nicht unterbrechen zu müssen.

§ 76. Fortsetzung. Die Hellingersche Zerlegung.

Ist kein Streckenspektrum vorhanden, so daß jedes $\tilde{\sigma}_{ik}^*(\mu)$ für alle μ verschwindet, so wird die Reduktion von $\|\sigma_{ik}(\mu)\|$ auf eine Normalform bereits durch (269) geleistet. Es möge also $\|\sigma_{ik}(\mu)\|$ ein Streckenspektrum haben. Bezeichnet a ein Intervall, so ist die Matrix $\|\underset{\mathrm{a}}{\Delta}\tilde{\sigma}_{ik}^*\|$, wie wir wissen, eine Einzelmatrix, also nichtnegativ definit, so daß also auch die binäre Hermitesche Kopplungsform

$$\Phi(\|\underset{\mathrm{a}}{\Delta}\tilde{\sigma}_{ik}^*\|;\ u+v x_1,\ v x_2,\ v x_3, \ldots) \equiv A u \bar{u} + \overline{B} u \bar{v} + B \bar{u} v + C v \bar{v}$$

nichtnegativ definit, mithin nach S. 45 die Determinante

$$AC - |B|^2 \geqq 0; \qquad A \geqq 0,\ C \geqq 0,$$

ist, d. h. es gilt

$$(180)' \qquad \Big|\sum_{k=1}^{\infty} \underset{\mathrm{a}}{\Delta}\tilde{\sigma}_{1k}^*(\mu) x_k\Big|^2 \leqq \Phi(\|\underset{\mathrm{a}}{\Delta}\tilde{\sigma}_{ik}^*\|;\ x_1, x_2, \ldots)\ \underset{\mathrm{a}}{\Delta}\tilde{\sigma}_{11}^*.$$

Da offenbar $\Phi(\|\underset{\mathrm{a}}{\Delta}\|) = \underset{\mathrm{a}}{\Delta}\Phi \leqq 1$ gilt, wenn $\mathfrak{x}$ auf E liegt, so können

wir mit Rücksicht auf (180) die Hellingerschen Integrale

$$(270)\qquad 0 \leqq \int_{-\infty}^{\mu} \frac{\left| d \sum\limits_{k=1}^{\infty} \tilde{\sigma}_{1k}^{*}(\mu)\, x_k \right|^2}{d\tilde{\sigma}_{11}^{*}(\mu)} \leqq 1$$

bilden, also auch die Matrix

$$(271)\qquad \| \tau_{ik}^{(1)}(\mu) \| = \left\| \int_{-\infty}^{\mu} \frac{d\tilde{\sigma}_{1i}^{*}(\mu)\, d\bar{\tilde{\sigma}}_{1k}^{*}(\mu)}{d\tilde{\sigma}_{11}^{*}(\mu)} \right\|,$$

deren Kopplungsform unter (270) steht. Setzen wir

$$(272)\qquad \| \sigma_{ik}^{(1)}(\mu) \| = \| \tilde{\sigma}_{ik}^{*}(\mu) - \tau_{ik}^{(1)}(\mu) \|,$$

so ist wegen $\tau_{1k}^{(1)} = \bar{\tau}_{k1}^{(1)} = \tilde{\sigma}_{1k}^{*}$

$$(273)\qquad \sigma_{1k}^{(1)}(\mu) \equiv \sigma_{k1}^{(1)}(\mu) \equiv 0; \qquad k = 1, 2, \ldots .$$

Man erkennt ohne Mühe, daß $\| \sigma_{ik}^{(1)}(\mu) \|$ eine spektrale Einzelmatrix von durchweg stetigen Funktionen ist. Wir können daher, entsprechend (271), die Einzelmatrix

$$(274)\qquad \| \tau_{ik}^{(2)}(\mu) \| = \left\| \int_{-\infty}^{\mu} \frac{d\sigma_{2i}^{(1)}(\mu)\, d\bar{\sigma}_{2k}^{(1)}(\mu)}{d\sigma_{22}^{(1)}(\mu)} \right\|$$

einführen,

$$(275)\qquad \| \sigma_{ik}^{(2)}(\mu) \| = \| \sigma_{ik}^{(1)} - \tau_{ik}^{(2)}(\mu) \|$$

setzen und, entsprechend (273), finden, daß in $\| \sigma_{ik}^{(2)}(\mu) \|$ alle Elemente der beiden ersten Zeilen und Kolonnen verschwinden. Allgemein setzen wir

$$(276)\qquad \begin{gathered} \| \sigma_{ik}^{(n)}(\mu) \| = \| \sigma_{ik}^{(n-1)}(\mu) - \tau_{ik}^{(n)}(\mu) \|, \\ \| \tau_{ik}^{(n)}(\mu) \| = \left\| \int_{-\infty}^{\mu} \frac{d\sigma_{ni}^{(n-1)}(\mu)\, d\bar{\sigma}_{nk}^{(n-1)}(\mu)}{d\sigma_{nn}^{(n-1)}(\mu)} \right\|, \\ n = 1, 2, \ldots; \qquad \sigma_{ik}^{(0)} = \tilde{\sigma}_{ik}^{*}. \end{gathered}$$

Dann ist $\| \sigma_{ik}^{(n)}(\mu) \|$ eine spektrale Einzelmatrix (von stetigen Funktionen) derart, daß die Elemente der n ersten Zeilen und Kolonnen gewiß verschwinden. Man hat

$$(277)\qquad \| \tilde{\sigma}_{ik}^{*}(\mu) \| = \sum_{j=1}^{n} \| \tau_{ik}^{(j)}(\mu) \| + \| \sigma_{ik}^{(n)}(\mu) \|,$$

und die $\| \tau_{ik}^{(j)}(\mu) \|$ sind, wie man ebenso wie bei der Jacobischen Transformation auf S. 60 verifizieren kann, paarweise orthogonal, unabhängig davon, ob in dem Paare dieselben oder zwei verschiedene Werte von μ betrachtet werden. Außerdem ist jedes $\| \tau_{ik}^{(j)}(\mu) \|$ eine spektrale Einzelmatrix, also ist die Reihe $\sum\limits_{j=1}^{\infty} \| \tau_{ik}^{(j)}(\mu) \|$ nach S. 158 konvergent und zwar $= \tilde{\sigma}_{ik}^{*}(\mu)$, da in dem Restglied $\| \sigma_{ik}^{(n)}(\mu) \|$ die n ersten Zeilen und Kolonnen wegen (276), (277) gleich Null sind. Man kann daher in der Zerlegung

$$(278) \qquad \| \tilde{\sigma}_{ik}^{*}(\mu) \| = \sum_{j=1}^{\infty} \| \tau_{ik}^{(j)}(\mu) \|$$

die Summanden gemäß dem Hauptsatze über Einzelmatrizen zunächst bei festem μ trennen und sodann die Schlußweise von S. 158 sinngemäß wiederholen. — Dies ist die Hellingersche Zerlegung der spektralen Einzelmatrizen (von stetigen Funktionen). Wir mußten uns dabei etwas kurz fassen und verweisen des näheren auf die Dissertation von Hellinger. Wir müssen leider auch darauf verzichten, den so gewonnenen Apparat in Verbindung mit § 50 auf das in der Hellingerschen Dissertation gelöste Äquivalenzproblem anzuwenden.

Fünftes Kapitel.

Spektraltheorie der beschränkten Matrizen.

§ 77. Das Auswahlverfahren.

Es sei $\mathfrak{A} = \| a_{ik} \|$ eine beschränkte Hermitesche Matrix, es seien $\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_n^{(n)}$ die Eigenwerte des n-ten Abschnittes $\mathfrak{A}_{[n]} = \| a_{ik} \|_{[n]}$, ferner sei $\| R_{ik}^{(n)}(\lambda) \|_{(n)}$ die Resolvente von $\mathfrak{A}_{[n]}$, also

$$(279) \qquad R_{ik}^{(n)}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\sigma_{ik}^{(n)}(\mu)}{\lambda - \mu}; \qquad i, k = 1, 2, \ldots, n,$$

wobei $\| \sigma_{ik}^{(n)}(\mu) \|$ die Spektralmatrix von $\mathfrak{A}_{[n]}$ bezeichnet. Ist Ω ein Gebiet in der λ-Ebene, dessen sämtliche Punkte von der reellen Achse oder wenigstens von denjenigen Punkten der reellen Achse, die dem Abschnittsspektrum angehören, in einer Entfernung liegt, die für alle Punkte von Ω nicht kleiner ist als eine feste Zahl $\varepsilon > 0$, so gilt auf Ω die Ungleichheit

$$(280) \qquad | R_{ik}^{(n)}(\lambda) | \leqq \varepsilon^{-1},$$

da $| \lambda - \mu | \leqq \varepsilon$, und, wie wir wissen, $\int_{-\infty}^{+\infty} | d\sigma_{ik}^{(n)}(\mu) | \leqq 1$ gilt [vgl. S. 51 sowie die Ungleichung (143) auf S. 85]. Die Kopplungsform

$$(281) \qquad \Phi(\| R_{ik}^{(n)}(\lambda) \|_{(n)}; x_1, x_2, \ldots, x_n) = \int_{-\infty}^{+\infty} \frac{1}{\lambda - \mu} \, d \sum_{i=1}^{n} \sum_{k=1}^{n} \sigma_{ik}^{(n)}(\mu) \, x_i \bar{x}_k$$

ist auf $\mathsf{E}_{[n]}$ und Ω absolut ebenfalls $\leqq \varepsilon^{-1}$, da die nicht abnehmende Funktion

$$(282) \qquad \xi^{(n)}(\mu; x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \sum_{k=1}^{n} \sigma_{ik}^{(n)}(\mu) \, x_i \bar{x}_k$$

auf $\mathsf{E}_{[n]}$ nicht negativ und $\leqq 1$ ist. Nach dem Satze von Montel gibt es

also zunächst eine Teilfolge $\{h_n\}$ von $\{n\}$ derart, daß $\lim\limits_{n=\infty} R_{ik}^{(h_n)}(\lambda)$ auf Ω vorhanden ist, wobei wir mit Rücksicht auf das Cantorsche Diagonalprinzip annehmen können, daß $\{h_n\}$ von i und k nicht abhängt. Es bezeichne $R_{ik}(\lambda)$ diesen Grenzwert. Dann ist $\mathsf{M}(\| R_{ik}(\lambda) \|)$ wegen (280) offenbar $\leqq \varepsilon^{-1}$, mithin gilt nach (233), (235), (212)

$$(283) \quad \Big|\sum_{k=1}^{\infty} R_{ik}(\lambda) x_k\Big|^2 \leqq \sum_{i=1}^{\infty} \Big|\sum_{k=1}^{\infty} R_{ik}(\lambda) x_k\Big|^2 = \Phi(\| R_{ik}(\lambda) \| \; \| R_{ik}(\lambda) \|^*)$$

$$\leqq \mathsf{M}(\| R_{ik}(\lambda) \| \; \| R_{ik}(\lambda) \|^*) = \mathsf{P}(\| R_{ik}(\lambda) \|)^2 \leqq 4\,\mathsf{M}(\| R_{ik}(\lambda) \|)^2 \leqq 4\varepsilon^{-2},$$

also nach der Umkehrung der Schwarzschen Ungleichung offenbar

$$\sum_{k=1}^{\infty} | R_{ik}(\lambda) |^2 \leqq 4\varepsilon^{-2}$$

und ebenso

$$\sum_{k=1}^{m} | R_{ik}^{(m)}(\lambda) |^2 \leqq 4\varepsilon^{-2},$$

d. h.

$$\sum_{k=1}^{\infty} | R_{ik}^{(m)}(\lambda) |^2 \leqq 4\varepsilon^{-2},$$

wobei $R_{ik}^{(m)}(\lambda) = 0$ gesetzt ist, wenn mindestens einer der beiden unteren Zeiger $> m$ ausfällt. Nach dem Majorantenkriterium der starken Konvergenz gilt also wegen $R_{ik}^{(h_n)} \to R_{ik}$

$$(284) \qquad \lim_{n=\infty} \sum_{k=1}^{h_n} c_k R_{ik}^{(h_n)}(\lambda) = \sum_{k=1}^{\infty} c_k R_{ik}(\lambda),$$

wenn $\{c_k\}$ eine quadratisch konvergente Zahlenfolge ist. Bezeichnet man den Grenzwert (284) mit X_i, so ist $\sum\limits_{i=1}^{\infty} | X_i |^2$ wegen (283) offenbar

$$\leqq 4\varepsilon^{-2} \sum_{k=1}^{\infty} | c_k |^2,$$

d. h. $\{X_i\}$ ist quadratisch konvergent. — Da die Wachstumstellen der Treppenfunktionen $\sigma_{ik}^{(n)}(\mu)$ die Abschnittseigenwerte darstellen, und diese durchweg reell sind, so ist auf Ω das Abschnittssystem

$$(285) \qquad \lambda x_i^{(h_n)} - \sum_{k=1}^{h_n} a_{ik} x_k^{(h_n)} = c_i; \qquad i = 1, 2, \ldots, h_n$$

gewiß lösbar, und zwar ist $x_i^{(h_n)}$, nach der Definition der Resolvente, die in (284) linkerhand hinter dem Limeszeichen stehende endliche

Summe. Es gilt dabei $\mathsf{M}(\| R_{ik}^{(m)}(\lambda) \|) \leqq \varepsilon^{-1}$, also wieder

$$\sum_{i=1}^{m} | x_i^{(m)} |^2 \leqq 4\varepsilon^{-2} \sum_{k=1}^{\infty} | c_k |^2.$$

Endlich ist $\mathfrak{A}$ eine Q-Matrix, so daß wegen $x_i^{(h_n)} \to X_i$ und mit Rücksicht auf das Majorantenkriterium analog zu (284) offenbar

$$\sum_{k=1}^{h_n} a_{ik} x_k^{(h_n)} \to \sum_{k=1}^{\infty} a_{ik} X_k,$$

also wegen (284), (285) auch

$$\lambda X_i - \sum_{k=1}^{\infty} a_{ik} X_k = c_i$$

gilt, so daß, wegen

$$\sum_{i=1}^{\infty} | X_i |^2 < +\infty$$

und mit Rücksicht auf (283), die zu $\lambda \mathfrak{E} - \mathfrak{A}$ gehörigen Gleichungen für jede quadratisch konvergente rechte Seite die quadratisch konvergente Lösung

$$X_i = \sum_{k=1}^{\infty} c_k R_{ik}(\lambda) \tag{286}$$

zulassen, und $\| R_{ik}(\lambda) \|$ ist beschränkt.

§ 78. Die Überflüssigkeit des Auswahlverfahrens.

Da $\mathfrak{A}$ beschränkt und von Hermiteschem Typus ist, könnte daraus nach Kap. III sofort geschlossen werden, daß $\| R_{ik}(\lambda) \|$ eine hintere und zugleich auch vordere Reziproke von $\lambda \mathfrak{E} - \mathfrak{A}$ darstellt. Damit die Schlußweise auch für allgemeinere, in Kap. VI zu betrachtende Matrizen in Kraft bleibt, bemerken wir, daß dies auch sonst erschlossen werden kann. $\lambda \mathfrak{E}_{[n]} - \mathfrak{A}_{[h_n]}$ hat, als eine endliche Matrix, nur eine Reziproke, also ist $\| R_{ik}^{(h_n)}(\lambda) \|$ vordere und hintere Reziproke von $\lambda \mathfrak{E}_{[h_n]} - \mathfrak{A}_{[h_n]}$, d. h. es gilt

$$\sum_{j=1}^{h_n} (\lambda \delta_{ij} - a_{ij}) R_{jk}^{(h_n)}(\lambda) = \sum_{j=1}^{h_n} R_{ij}^{(h_n)}(\lambda) (\lambda \delta_{jk} - a_{jk}) = \delta_{ik}. \tag{287}$$

Nimmt man hierbei den Grenzübergang $n \to +\infty$ unter Beachtung des Umstandes, daß $\mathfrak{A}$ eine Q-Matrix ist, und unter Benutzung des Majorantenkriteriums vor, so folgt

$$\sum_{j=1}^{\infty} (\lambda \delta_{ij} - a_{ij}) R_{jk}(\lambda) = \sum_{j=1}^{\infty} R_{ij}(\lambda) (\lambda \delta_{jk} - a_{jk}) = \delta_{ik}, \tag{288}$$

d. h., wie behauptet,

$$(289)\qquad (\lambda\mathfrak{E}-\mathfrak{A})\,\|R_{ik}(\lambda)\| = \|R_{ik}(\lambda)\|\,(\lambda\mathfrak{E}-\mathfrak{A}) = \mathfrak{E}.$$

Dabei ist in dem vorliegenden Fall das Auswahlverfahren überflüssig. Denn wäre $\lim\limits_{n=\infty} R_{ik}^{(n)}(\lambda)$ bei mindestens einem Wertpaare (i_0, k_0) nicht vorhanden, so würde es nach S. 83 zwei Auswahlfolgen geben, die bei diesem (i_0, k_0) zwei verschiedene Grenzwerte, etwa $R_{i_0k_0}^{*}(\lambda)$ und $R_{i_0k_0}^{**}(\lambda)$ liefern. Indem man jede der beiden Teilfolgen demselben Auswahlverfahren wie auf S. 81 die Folge $\{n\}$ unterwirft, gewinnt man die Existenz von zwei quadratisch konvergenten Lösungen

$$X_i^{*} = \sum_{k=1}^{\infty} R_{ik}^{*}(\lambda)\,c_k, \qquad X_i^{**} = \sum_{k=1}^{\infty} R_{ik}^{**}(\lambda)\,c_k,$$

die wegen $R_{i_0k_0}^{*}(\lambda) \neq R_{i_0k_0}^{**}(\lambda)$ nicht für alle quadratisch konvergenten Wertsysteme $\{c_i\}$ zusammenfallen können. Dies ist aber unmöglich, da $\lambda\mathfrak{E}-\mathfrak{A}$ eine beschränkte und normale Matrix ist, also nach § 60 nur eine beschränkte Reziproke haben kann.

Mithin ist das Auswahlverfahren überflüssig, es gilt $R_{ik}^{(n)}(\lambda) \to R_{ik}(\lambda)$ für $n \to \infty$, also auch $\sum\limits_{i=1}^{m}\sum\limits_{k=1}^{m} R_{ik}^{(n)}(\lambda)\,x_i\bar{x}_k \to \sum\limits_{i=1}^{m}\sum\limits_{k=1}^{m} R_{ik}(\lambda)\,x_i\bar{x}_k$ für ein jedes m. Der Ausdruck rechterhand ist aber, mit der auf S. 169 eingeführten Bezeichnung, gleich

$$(290)\qquad \int_{-\infty}^{+\infty} \frac{d\,\xi^{(n)}(\mu; x_1, \ldots, x_m, 0, \ldots, 0)}{\lambda-\mu},$$

wobei die $n-m$ letzten x gleich Null gesetzt werden. Da die ξ nach S. 169 nicht abnehmende, für $\mu < \mathfrak{n}(\mathfrak{A}) \leqq \mathfrak{n}(\mathfrak{A}_{[n]})$ identisch verschwindende Funktionen darstellen und für $\sum\limits_{i=1}^{\infty} |x_i|^2 \leqq 1$ durchweg $\leqq 1$ sind, so folgt nach dem Grommer-Hamburgerschen Kriterium (S. 105), daß $\xi^{(n)}$ für $n \to +\infty$ „im wesentlichen" gegen eine Grenzfunktion ξ und das Integral (290) gegen

$$\int_{-\infty}^{+\infty} \frac{d\xi(\mu; x_1, x_2, \ldots, x_m, 0, \ldots, 0)}{\lambda-\mu}$$

konvergiert. Und zwar ist ξ offenbar von der Gestalt

$$\xi(\mu; x_1, \ldots, x_m, 0, \ldots, 0) = \sum_{i=1}^{m}\sum_{k=1}^{m} \sigma_{ik}(\mu)\,x_i\bar{x}_k,$$

es gilt daher

$$\sum_{i=1}^{m}\sum_{k=1}^{m}\sigma_{ik}^{(n)}(\mu)\,x_i\bar{x}_k\ [\rightarrow]\ \sum_{i=1}^{m}\sum_{k=1}^{m}\sigma_{ik}(\mu)\,x_i\bar{x}_k;\qquad n\rightarrow+\infty.$$

Hierbei ist m eine feste, jedoch beliebig große Zahl, und die x_i sind beliebig. Setzt man $x_i=1$ und die übrigen $x=0$, so erhält man $\sigma_{ii}^{(n)}(\mu)\,[\rightarrow]\,\sigma_{ii}(\mu)$. Ist also $i\neq k$ und setzt man alle x außer x_i und x_k gleich Null, so findet man wegen $\sigma_{ik}^{(n)}=\bar{\sigma}_{ki}^{(n)}$

$$\sigma_{ik}^{(n)}(\mu)\,x_i\bar{x}_k+\bar{\sigma}_{ik}^{(n)}(\mu)\,\bar{x}_i x_k\ [\rightarrow]\ \sigma_{ik}(\mu)\,x_i\bar{x}_k+\bar{\sigma}_{ik}(\mu)\,\bar{x}_i x_k.$$

Setzt man $x_i=1$, $x_k=1$, sodann $x_i=1$, $x_k=\sqrt{-1}$, so folgt

$$\sigma_{ik}^{(n)}(\mu)\pm\bar{\sigma}_{ik}^{(n)}(\mu)\ [\rightarrow]\ \sigma_{ik}(\mu)\pm\bar{\sigma}_{ik}(\mu),$$

also offenbar $\sigma_{ik}^{(n)}(\mu)\,[\rightarrow]\,\sigma_{ik}(\mu)$, was somit auch für $i\neq k$ gilt. — Nimmt man in dem bei (279), (286) Gesagten dieselbe Spezialisierung der x vor, so folgt ebenso

$$(291)\qquad \int_{-\infty}^{+\infty}\frac{d\sigma_{ik}^{(n)}(\mu)}{\lambda-\mu}=R_{ik}^{(n)}(\lambda)\rightarrow R_{ik}(\lambda)=\int_{-\infty}^{+\infty}\frac{d\sigma_{ik}(\mu)}{\lambda-\mu};\qquad i=1,2,\ldots$$

für alle λ in Ω. Alle beliebigen nicht reellen λ sowie diejenigen reellen λ, die nicht Häufungsstellen der Abschnittseigenwerte sind, also gewiß alle reellen Zahlen, die $<\mathsf{n}(\mathfrak{A})\leqq\mathsf{n}(\mathfrak{A}_{[n]})$ oder $>\mathsf{m}(\mathfrak{A})\geqq\mathsf{m}(\mathfrak{A}_{[n]})$ sind, müssen einem Gebiete Ω angehören. Für die zuletzt genannten μ ist nach S. 51 $\sigma_{ik}^{(n)}(\mu)=0$ bzw. $\sigma_{ik}^{(n)}(\mu)=\delta_{ik}$, so daß auch

$$(292)\qquad \sigma_{ik}(\mu)\equiv 0,\quad \mu<\mathsf{n}(\mathfrak{A});\qquad \sigma_{ik}(\mu)\equiv\delta_{ik},\quad \mu>\mathsf{m}(\mathfrak{A})$$

gilt und daher

$$(293)\qquad R_{ik}(\lambda)=\int_{-\infty}^{+\infty}\frac{d\sigma_{ik}(\mu)}{\lambda-\mu}=\int_{\mathsf{n}(\mathfrak{A})-0}^{\mathsf{m}(\mathfrak{A})+0}\frac{d\sigma_{ik}(\mu)}{\lambda-\mu}$$

ist, und die $\sigma_{ik}(\mu)$ durch die weitere Forderung der rechtsseitigen Stetigkeit eindeutig bestimmt sind. Die hierdurch festgelegte unendliche Hermitesche Matrix $\|\sigma_{ik}(\mu)\|$ nennen wir Spektralmatrix von $\mathfrak{A}$; ihre Kopplungsform ist in jedem Punkte von E eine nicht abnehmende, für $\mu<\mathsf{n}(\mathfrak{A})$ verschwindende Funktion von μ, die für $\mu>\mathsf{m}(\mathfrak{A})$ gleich 1 ist. Insbesondere sind also die $\|\sigma_{ik}(\mu)\|$ gleichmäßig beschränkt. Letzteres wird auch gelten, wenn $\mathfrak{A}$ nicht beschränkt ist (vgl. Kap. VI), da entsprechendes für alle Abschnittsmatrizen gilt (Kap. I).

§ 79. Die Spektralmatrix als spektrale Einzelmatrix.

Wir beweisen nun, daß die Spektralmatrix jeder beschränkten Hermiteschen Matrix eine spektrale Einzelmatrix ist. Da die $\sigma_{ik}(\mu)$ für $\mu < \mathsf{n}(\mathfrak{A})$ verschwinden, genügt es, wenn wir anstatt von (258) die Relation

$$\left\| \underset{\mathrm{a}}{\Delta} \sigma_{ik} \right\| \left\| \underset{\mathrm{b}}{\Delta} \sigma_{ik} \right\| = \left\| \underset{\mathrm{ab}}{\Delta} \sigma_{ik} \right\| \tag{294}$$

beweisen. Auch (294) brauchen wir nur für den Fall zu beweisen, wo sowohl a als auch b ein Punkt, etwa μ_1 bzw. μ_2 ist. Man erkennt dies leicht, wenn man wieder beachtet, daß $\sigma_{ik}(\mu) \equiv 0$ gilt für $\mu < \mathsf{n}(\mathfrak{A})$. Wir werden dabei genau so wie auf S. 56 vorzugehen haben; die durch (81) in Evidenz gesetzte Übereinstimmung der Residua wird dabei durch die Hilbertsche „Residuenformel“ (169) zu ersetzen sein.

Wir gehen also wieder von der Hilbertschen Funktionalgleichung (S. 143), d. h. der Gleichung

$$(\lambda_1 - \lambda_2) \| R_{ik}(\lambda_1) \| \, \| R_{ik}(\lambda_2) \| = \| R_{ik}(\lambda_2) - R_{ik}(\lambda_1) \| \tag{295}$$

aus, wobei λ_1 und λ_2 zwei etwa in der oberen Halbebene gelegene Zahlen bezeichnen mögen; dann sind λ_1 und λ_2 gewiß reguläre Stellen von $\mathfrak{A}$ (da, wie wir gesehen haben, eine beschränkte Reziproke für $\lambda \mathfrak{E} - \mathfrak{A}$ vorhanden ist), so daß nach S. 143 die Hilbertsche Funktionalgleichung gewiß gilt. Aus (293) und (295) folgt, wenn $\lambda_\nu = \mu_\nu + \varepsilon_\nu \sqrt{-1}$, $\varepsilon_\nu > 0$, $\mu_\nu \lesseqgtr 0$, $\nu = 1, 2$ gesetzt wird,

$$[\mu_1 - \mu_2 + \sqrt{-1}(\varepsilon_1 - \varepsilon_2)] \sum_{j=1}^{\infty} \int_{-\infty}^{+\infty} \frac{d\sigma_{ij}(\mu)}{\mu_1 + \varepsilon_1\sqrt{-1} - \mu} \int_{-\infty}^{+\infty} \frac{d\sigma_{jk}(\mu)}{\mu_2 + \varepsilon_2\sqrt{-1} - \mu}$$

$$= -\int_{-\infty}^{+\infty} \frac{d\sigma_{ik}(\mu)}{\mu_1 + \varepsilon_1\sqrt{-1} - \mu} + \int_{-\infty}^{+\infty} \frac{d\sigma_{ik}(\mu)}{\mu_2 + \varepsilon_2\sqrt{-1} - \mu}. \tag{296}$$

Wir multiplizieren hier beiderseits mit $\varepsilon_1 \sqrt{-1}$ und nehmen bei festem $\varepsilon_2 > 0$ den Grenzübergang $\varepsilon_1 \to +0$ vor. Der Grenzwert der rechten Seite ist nach der „Residuenformel“ (§ 44) gleich $-\underset{\mu_1}{\Delta} \sigma_{ik}$. Die linke Seite von (296) wird nach der Multiplikation mit $\varepsilon_1 \sqrt{-1}$ mit Rücksicht auf (293) gleich

$$\sqrt{-1}\,\varepsilon_1 \int_{-\infty}^{+\infty} \frac{d \sum_{j=1}^{\infty} \sigma_{ij}(\mu) \cdot R_{jk}(\mu_2 + \varepsilon_2\sqrt{-1}) \cdot (\mu_1 - \mu_2 + \sqrt{-1}\,\varepsilon_1 - \sqrt{-1}\,\varepsilon_2)}{\mu_1 + \varepsilon_1\sqrt{-1} - \mu}. \tag{297}$$

[Die vorgenommene Vertauschung von $\int\limits_{-\infty}^{+\infty}$ mit $\sum\limits_{j=1}^{\infty}$ rechtfertigt sich dabei wie folgt. Die Matrix $\| R_{jk}(\mu_2 + \varepsilon_2 \sqrt{-1}) \|$ ist beschränkt, also ist $\sum\limits_{j=1}^{\infty} | R_{jk}(\mu_2 + \varepsilon_2 \sqrt{-1}) |^2$, bei festem k konvergent; daher ist nach S. 127 die Reihe $\sum\limits_{j=1}^{\infty} x_j R_{jk}(\mu_2 + \varepsilon_2 \sqrt{-1})$ für $\sum\limits_{j=1}^{\infty} | x_j |^2 \leqq 1$ gleichmäßig konvergent. Nun gilt aber

$$\mathsf{M}(\| \sigma_{ik}(\mu) \|) \leqq 1, \quad \text{also} \quad \sum_{j=1}^{\infty} | \sigma_{ij}(\mu) |^2 \leqq 1,$$

mithin ist

$$\sum_{j=1}^{\infty} \sigma_{ij}(\mu) R_{jk}(\mu_2 + \varepsilon_2 \sqrt{-1})$$

für alle μ gleichmäßig konvergent. Beachtet man daher, daß die σ_{ik} wegen (292) außerhalb eines hinreichend großen Intervalles konstant sind, so daß der Integrationsweg mit Rücksicht auf (293) eigentlich endlich ist, so erkennt man, daß der Satz von § 41 angewandt werden kann. Die in (297) hinter dem Zeichen der Integration stehende Funktion von μ ist nämlich von beschränkter Schwankung. Es genügt zu zeigen, daß die Funktion

$$\sum_{j=1}^{\infty} \sigma_{ij}(\mu) \eta_j$$

von μ von beschränkter Schwankung ist, wenn

$$\sum_{j=1}^{\infty} | \eta_j |^2 < +\infty$$

gilt. Dies folgt aber daraus, daß

$$\sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \sigma_{ik}(\mu) x_i \bar{x}_k$$

eine Hermitesche Form ist, die für $\mu < \mathsf{n}(\mathfrak{A})$ verschwindet, für $\mu > \mathsf{m}(\mathfrak{A})$ gleich

$$\sum_{j=1}^{\infty} | x_j |^2$$

wird und nie abnimmt, also von beschränkter Schwankung ist. Setzt man in der Tat $x_1 = 0$ und läßt man die übrigen x willkürlich, so folgt zunächst, daß auch

$$\sum_{i=2}^{\infty} \sum_{k=2}^{\infty} \sigma_{ik}(\mu) x_i \bar{x}_k$$

von beschränkter Schwankung ist, also ist es auch die Differenz der

beiden Doppelsummen, mithin, in Anbetracht des Hermiteschen Charakters, auch

$$\sum_{j=2}^{\infty} \sigma_{1j}(\mu)\, \eta_j,$$

während $\sigma_{11}(\mu)$ allein gewiß von beschränkter Schwankung ist, da nach S. 51 jedes $\sigma_{ik}^{(n)}(\mu)$ eine Totalschwankung $\leqq 1$ hat. Damit ist die Behauptung für $i = 1$ bewiesen. Die übrigen i können offenbar ebenso erledigt werden. — Wir werden diese naheliegenden Betrachtungen später nicht mehr wiederholen.]

Nimmt man in (297) den Grenzübergang $\varepsilon_1 \to +0$ vor, so findet man nach der „Residuenformel" (S. 97)

$$\sum_{j=1}^{\infty} \underset{\mu_1}{\Delta} \sigma_{ij} \cdot R_{jk}(\mu_2 + \varepsilon_2 \sqrt{-1}) \cdot (\mu_1 - \mu_2 - \sqrt{-1}\,\varepsilon_2). \tag{298}$$

Dies ist der Grenzwert der mit $\varepsilon_1 \sqrt{-1}$ multiplizierten linken Seite von (296), während wir rechterhand $-\underset{\mu_1}{\Delta} \sigma_{ik}$ gefunden haben. Mithin ist

$$(\mu_2 - \mu_1 + \varepsilon_2 \sqrt{-1}) \sum_{j=1}^{\infty} \underset{\mu_1}{\Delta} \sigma_{ij} \int_{-\infty}^{+\infty} \frac{d\sigma_{jk}(\mu)}{\mu_2 + \varepsilon_2 \sqrt{-1} - \mu} = \underset{\mu_1}{\Delta} \sigma_{ik}. \tag{299}$$

Der Grenzübergang $\varepsilon_2 \to +0$ ergibt daraus wegen (169), wenn zunächst $\mu_1 = \mu_2$ angenommen wird,

$$\sum_{j=1}^{\infty} \underset{\mu_2}{\Delta} \sigma_{ij} \underset{\mu_1}{\Delta} \sigma_{jk} = \underset{\mu_1}{\Delta} \sigma_{ik}; \qquad \mu_1 = \mu_2,$$

so daß die zu beweisende Beziehung

$$\| \underset{\mu_1}{\Delta} \sigma_{ik} \| \; \| \underset{\mu_2}{\Delta} \sigma_{ik} \| = \gamma \, \| \underset{\mu_1}{\Delta} \sigma_{ik} \|; \qquad \begin{array}{ll} \gamma = 0 & \text{für } \mu_1 \neq \mu_2, \\ \gamma = 1 & \text{für } \mu_1 = \mu_2 \end{array} \tag{300}$$

für $\mu_1 = \mu_2$ gewiß richtig ist. Es sei also $\mu_1 \neq \mu_2$. Multiplizieren wir (299) beiderseits mit $\varepsilon_2 \sqrt{-1} \to 0$, so kommt rechterhand Null, linkerhand aber offenbar nur

$$\lim_{\varepsilon_2 \to +0} \varepsilon_2 \sqrt{-1}\, (\mu_1 - \mu_2) \sum_{j=1}^{\infty} \underset{\mu_1}{\Delta} \sigma_{ij} \int_{-\infty}^{+\infty} \frac{d\sigma_{jk}(\mu)}{\mu_2 + \varepsilon_2 \sqrt{-1} - \mu};$$

was aber nach der Residuenformel gleich

$$(\mu_1 - \mu_2) \sum_{j=1}^{\infty} \underset{\mu_2}{\Delta} \sigma_{ij} \underset{\mu_1}{\Delta} \sigma_{jk}$$

ist. Wegen $\mu_1 \neq \mu_2$ ist also auch

$$\sum_{j=1}^{\infty} \underset{\mu_2}{\Delta} \sigma_{ij} \underset{\mu_1}{\Delta} \sigma_{jk} = 0,$$

so daß (300) auch für $\mu_1 \neq \mu_2$ gilt. — Damit ist gezeigt, daß die Spektralmatrix eine spektrale Einzelmatrix ist. — Auch bei den zwar nicht Hermiteschen, aber doch noch normalen beschränkten Matrizen scheint eine ähnliche Betrachtung möglich zu sein, da bei den normalen (und nur bei den normalen) Matrizen die beiden Hermiteschen Komponenten vertauschbar sind. Vgl. S. 33.

§ 80. Das Spektrum. Momentendarstellung der Matrizenpotenzen.

Das Spektrum der Spektralmatrix von $\mathfrak{A}$ wollen wir einfach als das Spektrum von $\mathfrak{A}$ bezeichnen. Ebenso sprechen wir von dem Streckenspektrum von $\mathfrak{A}$ usf. — Wir zeigen, daß, wenn λ nicht dem Spektrum von $\mathfrak{A}$ angehört, λ gewiß eine (in dem auf S. 142 erklärten matrizentheoretischen Sinne) reguläre Stelle von $\mathfrak{A}$ ist. Ist λ nicht reell, so ist dies selbstverständlich; ebenso, wenn $\lambda < \mathsf{n}(\mathfrak{A})$ oder $> \mathsf{m}(\mathfrak{A})$ gilt. Es handelt sich also nur um die eventuellen, und, wie wir wissen (S. 143), bei passendem $\mathfrak{A}$ immer möglichen „Spektrumslücken" zwischen $\mathsf{n}(\mathfrak{A})$ und $\mathsf{m}(\mathfrak{A})$. Es sei nun λ_0 eine reelle Zahl, die nicht dem Spektrum angehört. Dann gibt es also ein $2\varepsilon > 0$ derart, daß $\sigma_{ik}(\mu) \equiv \sigma_{ik}(\lambda_0)$ gilt für $\lambda_0 - 2\varepsilon < \mu < \lambda_0 + 2\varepsilon$. Auf diesem Intervall ist dann wegen (293) nach S. 144 jedes Element $R_{ik}(\lambda)$ der Resolvente von $\mathfrak{A}$ regulär. Die Kopplungsform von $\| R_{ik}(\lambda) \|$ ist in dem um den Punkt λ_0 mit dem Radius ε geschlagenen Kreise gleichmäßig beschränkt. Bezeichnet nämlich λ einen Punkt dieses Kreises K, so gilt, da in (293) die Differenz $|\mu - \lambda| \geqq \varepsilon$ angenommen werden kann, weil das Intervall $\lambda_0 - \varepsilon \leqq \mu \leqq \lambda_0 + \varepsilon$ ein Konstanzintervall aller $\sigma_{ik}(\mu)$ ist, offenbar

$$(301) \qquad |\Phi(\| R_{ik}(\lambda) \|_{[n]})| = \left| \int_{-\infty}^{+\infty} \frac{d \sum_{i=1}^{n} \sum_{k=1}^{n} \sigma_{ik}(\mu) x_i \bar{x}_k}{\lambda - \mu} \right|$$

$$\leqq \varepsilon^{-1} \int_{-\infty}^{+\infty} d \sum_{i=1}^{n} \sum_{k=1}^{n} \sigma_{ik}(\mu) x_i \bar{x}_k = \varepsilon^{-1} \sum_{i=1}^{n} |x_i|^2 \leqq \varepsilon^{-1},$$

da $\sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \sigma_{ik}(\mu) x_i \bar{x}_k$ eine nicht abnehmende Funktion von μ ist, die

für $\mu < \mathsf{n}(\mathfrak{A})$ gleich Null und für $\mu > \mathsf{m}(\mathfrak{A})$ gleich $\sum_{i=1}^{\infty} |x_i|^2$ ausfällt. — Mithin sind auch die Reihen $\sum_{k=1}^{\infty} |R_{ik}(\lambda)|^2$ auf K gleichmäßig beschränkt. Es gilt ferner, der Regularität von $R_{ik}(\lambda)$ zufolge, $R_{ik}(\lambda) \rightarrow R_{ik}(\lambda_0)$ für $\lambda \rightarrow \lambda_0$, wobei λ etwa in der oberen Halbebene liegen möge (während λ_0 reell ist). Bezeichnet also $\{c_k\}$ eine quadratisch konvergente Zahlenfolge, so gilt mit Rücksicht auf das Majorantenkriterium der starken Konvergenz die Grenzgleichung

$$\sum_{k=1}^{\infty} R_{ik}(\lambda)\, c_k \rightarrow \sum_{k=1}^{\infty} R_{ik}(\lambda_0)\, c_k, \tag{302}$$

und zwar ist die Reihe $\sum_{i=1}^{\infty} \left| \sum_{k=1}^{\infty} R_{ik}(\lambda_0)\, c_k \right|^2$ gewiß konvergent, da $\|R_{ik}(\lambda)\|$ auf K gleichmäßig beschränkt ist. Folglich ist $x_i = \sum_{k=1}^{\infty} c_k R_{ik}(\lambda_0)$ eine quadratisch konvergente Lösung des Gleichungssystems $(\lambda_0 \mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{c}$, indem $\sum_{k=1}^{\infty} c_k R_{ik}(\lambda)$ gewiß die Lösung von $(\lambda \mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{c}$ darstellt, da ja λ nicht reell ist. Da $\lambda_0 \mathfrak{E} - \mathfrak{A}$ Hermitesch ist, so ist der Fall **3** (S. 139) von vornherein ausgeschlossen, also liegt der Fall **1** vor, d. h. λ_0 ist eine reguläre Stelle der Matrix, w. z. b. w. Wir werden sogleich sehen, daß umgekehrt im Spektrum keine *beschränkte* Reziproke für $\mu \mathfrak{E} - \mathfrak{A}$ existieren kann. Das Spektrum ist also identisch mit der Menge der matrizentheoretisch singulären Stellen von $\mathfrak{A}$.

Die Punkte des Spektrums sind nach Definition Zahlen μ_0, denen kein von i und k unabhängiges $\varepsilon > 0$ zugeordnet werden können, so daß für $\mu_0 - \varepsilon \leqq \mu \leqq \mu_0 + \varepsilon$ ein jedes $\sigma_{ik}(\mu)$ konstant ist. Nach S. 98 und mit Rücksicht auf (293) kann es also kein $\varepsilon > 0$ geben derart, daß im Kreise $|\mu_0 - \lambda| < \varepsilon$ jedes $R_{ik}(\lambda)$ regulär ist. Mithin ist dann keine konvergente Entwicklung nach Potenzen von $\lambda - \mu_0$ möglich, folglich ist μ_0 nach S. 144 eine singuläre Stelle von $\mathfrak{A}$. Gehört umgekehrt μ_0 nicht dem Spektrum an, so ist μ_0 eine matrizentheoretisch reguläre Stelle, es gibt also nach S. 144 ein $\varepsilon > 0$ derart, daß im Kreise $|\mu_0 - \lambda| < \varepsilon$ die Taylorsche Matrizenreihe konvergiert und die Resolvente darstellt. Erst recht sind also die $R_{ik}(\lambda)$ für $|\mu_0 - \lambda| < \varepsilon$ regulär. Die matrizentheoretisch singulären Stellen einer beschränkten Hermiteschen Matrix sind also mit den singulären Stellen und den Häufungsstellen der singulären Stellen der Funktionen $R_{ik}(\lambda)$ identisch, wobei wir die Häufungsstellen der funktionentheoretisch singulären Stellen besonders erwähnen müssen, da i und k nicht fest sind.

Für hinreichend große $|\lambda|$ gilt nach S. 101 bzw. S. 143

$$(303)\qquad R_{ik}(\lambda)=\int\limits_{-\infty}^{+\infty}\frac{d\sigma_{ik}(\mu)}{\lambda-\mu}=\sum_{\nu=1}^{\infty}\frac{\int\limits_{-\infty}^{+\infty}\mu^{\nu}\,d\sigma_{ik}(\mu)}{\lambda^{\nu+1}}$$

und

$$(304)\qquad \|R_{ik}(\lambda)\|=(\lambda\mathfrak{E}-\mathfrak{A})^{-1}=\sum_{\nu=1}^{\infty}\frac{\mathfrak{A}^{\nu}}{\lambda^{\nu+1}},$$

folglich ist

$$(305)\qquad \mathfrak{A}^{\nu}=\Big\|\int\limits_{-\infty}^{+\infty}\mu^{\nu}\,d\sigma_{ik}(\mu)\Big\|;\qquad \nu=0,1,\ldots.$$

Ist $\lambda=0$ keine singuläre Stelle von $\mathfrak{A}$, so hat $-\mathfrak{A}$ (also auch $\mathfrak{A}$) eine beschränkte Reziproke und umgekehrt, und es ist dabei jedes $\sigma_{ik}(\mu)$ in einer Umgebung des Nullpunktes konstant. Für hinreichend kleine $|\lambda|$ gilt dann (vgl. S. 144)

$$(306)\qquad R_{ik}(\lambda)=\int\limits_{-\infty}^{+\infty}\frac{d\sigma_{ik}(\mu)}{\lambda-\mu}=-\sum_{\nu=0}^{\infty}\lambda^{\nu}\int\limits_{-\infty}^{+\infty}\frac{d\sigma_{ik}(\mu)}{\mu^{\nu+1}}$$

mit

$$(307)\qquad \|R_{ik}(\lambda)\|=(\lambda\mathfrak{E}-\mathfrak{A})^{-1}=-\sum_{\nu=0}^{\infty}\lambda^{\nu}\mathfrak{A}^{-\nu-1},$$

so daß (305) auch für $\nu=-1,-2,\ldots$ richtig ist.

§ 81. Über die Spektralmatrix der Reziproke u. dgl.

Wir wollen hier über die Spektralmatrix der Reziproke eine Bemerkung machen. Es sei zunächst $\mathfrak{A}$ eine beliebige beschränkte Hermitesche Matrix, die eine beschränkte Reziproke $\mathfrak{A}^{-1}$ hat, es sei $\|\sigma_{ik}(\mu)\|$ die Spektralmatrix von $\mathfrak{A}$ und $\|\tau_{ik}(\mu)\|$ diejenige von $\mathfrak{A}^{-1}$, endlich $\|S_{ik}(\lambda)\|$ die Resolvente von $\mathfrak{A}^{-1}$, also für hinreichend große $|\lambda|$

$$\|S_{ik}(\lambda)\|=\sum_{\nu=0}^{\infty}\frac{\mathfrak{A}^{-\nu}}{\lambda^{\nu+1}},\qquad \text{folglich}\qquad \mathfrak{A}^{-\nu}=\int\limits_{-\infty}^{+\infty}\frac{d\sigma_{ik}}{\mu^{\nu}}$$

(vgl. oben), mithin

$$(308)\qquad S_{ik}(\lambda)=\frac{1}{\lambda}\sum_{\nu=0}^{\infty}\int\limits_{-\infty}^{+\infty}\frac{d\sigma_{ik}(\mu)}{\lambda^{\nu}\mu^{\nu}}=\int\limits_{-\infty}^{+\infty}\frac{d\sigma_{ik}(\mu)}{\lambda-\dfrac{1}{\mu}},$$

wobei sich die gliedweise Integration genau so wie auf S. 175 rechtfertigt. Denn die Belegungen liegen diesmal nicht nur ganz im End-

lichen, sondern können auch dem Punkte $\mu = 0$ nicht beliebig nahe kommen. Die Integraldarstellung ist, aus Gründen der analytischen Fortsetzung, allgemein, d. h. auch außerhalb des Konvergenzgebietes der C. Neumannschen Reihen gültig. Andererseits haben wir

$$S_{ik}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\tau_{ik}(\mu)}{\lambda - \mu}. \tag{309}$$

Führen wir daher in (308) $\frac{1}{\mu}$ als Integrationsvariable ein, so kann $\tau_{ik}(\mu)$ aus (309) unmittelbar abgelesen werden (dank der durch die Stieltjessche Umkehrform gesicherten Eindeutigkeit). Soviel steht von vornherein fest, daß ein Punkt μ dann und nur dann dem Spektrum von $\mathfrak{A}$ angehört, wenn $\frac{1}{\mu}$ im Spektrum von $\mathfrak{A}^{-1}$ liegt. Es folgt nämlich aus den Sätzen von § 66, daß die beiden Hermiteschen Matrizen $\mu\mathfrak{E} - \mathfrak{A}$, $\frac{1}{\mu}\mathfrak{E} - \mathfrak{A}^{-1}$ nur zugleich eine beschränkte Reziproke haben können. Nehmen wir der Übersichtlichkeit halber an, daß das Spektrum von $\mathfrak{A}$ ganz auf der positiven Halbachse liegt. Dann trifft dies auch für das Spektrum von $\mathfrak{A}^{-1}$ zu, und $\frac{1}{\mu}$ ist auf den beiden Sprektren monoton. Für die Stieltjessche Integration kommt aber nur das Spektrum in Betracht, so daß wir so rechnen können, als ob $\frac{1}{\mu}$ von $-\infty$ bis $+\infty$ eine monotone, stetig differentiierbare Funktion usf. wäre. Denn $\mu = 0$ liegt außerhalb der Spektren. [Ebenso könnten wir den Fall betrachten, wo das Spektrum von $\mathfrak{A}$ auf der negativen Halbachse liegt. Eine der beiden Voraussetzungen trifft, wenn das Spektrum, wie oft, ein (zusammenhängendes) Intervall ist, von selbst zu, sobald nur $\mathfrak{A}^{-1}$ existiert.]

Ist $\sigma_{ik}(\mu)$ stetig differentiierbar, so gilt, wenn

$$\varrho_{ik}(\mu) = \int_{\frac{1}{\mu}}^{+\infty} \frac{d\sigma_{ik}(\nu)}{\nu^2} \tag{310}$$

gesetzt wird, die Umformung

$$S_{ik}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\sigma_{ik}(\mu)}{\lambda - \frac{1}{\mu}} = \int_{-\infty}^{+\infty} \frac{d\sigma_{ik}\left(\frac{1}{\mu}\right)}{\lambda - \mu} = \int_{-\infty}^{+\infty} \frac{d\varrho_{ik}(\mu)}{\lambda - \mu} \tag{308'}$$

gewiß. Doch ist die Umformung, wie durch Zurückgreifen auf die Definition der Stieltjesschen Integration unmittelbar folgt, auch sonst legal (vgl. nämlich § 40). Da σ_{ik} von rechts stetig ist, so ist ϱ_{ik} von links stetig. Offenbar gilt $\varrho_{ik}(+0) = 0$, und, wenn in (310) der passende „Zweig" gewählt wird, auch $\varrho_{ik}(\mu) = 0$, $\mu \leqq 0$. Aus (309) und (308)' folgt nun

$$(311) \qquad \tau_{ik}(\mu) \equiv \varrho_{ik}(\mu - 0).$$

Die Spektralmatrix $\| \tau_{ik} \|$ von $\mathfrak{A}^{-1}$ leitet sich also aus derjenigen von $\mathfrak{A}$ mittels (310), (311) ab. Diese Regel ist für die explizite Beherrschung der Division durch Fouriersche und verwandte Reihen von Wichtigkeit (vgl. S. 189ff.). Die Diagonalelemente der Spektralmatrix geben dabei die „statistische" Verteilung, d. h. die relative Häufigkeit der verschiedenen Teile des ganzen Wertgebietes der auf Fouriersche Weise entwickelten Funktion an. Wir kommen darauf bei den fastperiodischen Matrizen zurück (S. 251).

Es sei hier noch erwähnt, daß die Zurückführung der Spektralmatrix von $\varphi(\mathfrak{A})$ auf diejenige von $\mathfrak{A}$ nicht nur im vorliegenden Falle $\varphi(z) = \frac{1}{z}$, sondern auch bei anderen Funktionen von $\mathfrak{A}$ gelingt. In seiner Anwendung auf fastperiodische Matrizen gewährt dieses Prinzip, wie an anderer Stelle gezeigt werden soll, manche wesentliche (gruppentheoretische) Einblicke in die Natur von diophantischen Verteilungskopplungen und klärt so den wahren Grund für den von Bohr entdeckten Zusammenhang zwischen den Dirichletschen Reihen und den unendlich vielen Variablen auf (vgl. S. 46). — Wir müssen jedoch zunächst zu willkürlichen Funktionen von beliebigen beschränkten Hermiteschen Matrizen übergehen.

§ 82. Das Korrespondenzprinzip.

Aus (305) folgt

$$(312) \qquad \varphi(\mathfrak{A}) = \left\| \int_{-\infty}^{+\infty} \varphi(\mu)\, d\sigma_{ik}(\mu) \right\|,$$

wenn zunächst $\varphi(z)$ ein Polynom bezeichnet. Ist $\psi(z)$ ein anderes Polynom und $\omega(z)$ das Produkt $\varphi(z)\psi(z)$, so ist offenbar (vgl. S. 4) $\varphi(\mathfrak{A})\psi(\mathfrak{A}) = \omega(\mathfrak{A})$. Diese formale Bemerkung läßt sich wegen (312) sofort in die Gestalt der Hilbertschen Vollständigkeitsrelation der Spektralmatrix kleiden:

$$(313) \quad \left\| \int_{-\infty}^{+\infty} \varphi(\mu)\, d\sigma_{ik}(\mu) \right\| \left\| \int_{-\infty}^{+\infty} \psi(\mu)\, d\sigma_{ik}(\mu) \right\| = \left\| \int_{-\infty}^{+\infty} \varphi(\mu)\,\psi(\mu)\, d\sigma_{ik}(\mu) \right\|,$$

wobei φ und ψ zunächst nur Polynome sind.

Es sei $f(\mu)$ eine für alle μ, oder, was für unsere Zwecke dasselbe ist, für $\mathsf{n}(\mathfrak{A}) - 0 \leqq \mu \leqq \mathsf{m}(\mathfrak{A}) + 0$ erklärte, beschränkte und zunächst auch stetige Funktion. Die Matrix $\left\| \int\limits_{-\infty}^{+\infty} f(\mu)\, d\sigma_{ik}(\mu) \right\|$, die wir mit $f(\mathfrak{A})$ bezeichnen wollen, ist alsdann beschränkt. Bezeichnet nämlich M die obere Grenze von $f(\mu)$ und ist $\mu_1, \mu_2, \ldots, \mu_{N+1}$ irgendein System von Einteilungspunkten, so ist die Kopplungsform der Näherungsmatrix

$$\left\| \sum_{\nu=1}^{N} f(\mu_\nu) \left[\sigma_{ik}(\mu_{\nu+1}) - \sigma(\mu_\nu)\right] \right\|$$

absolut

$$\leqq M \sum_{\nu=1}^{N} \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} (\sigma_{ik}(\mu_{\nu+1}) - \sigma_{ik}(\mu_\nu))\, x_i \bar{x}_k \leqq M; \qquad \sum_{i=1}^{\infty} |x_i|^2 = 1,$$

da $\boldsymbol{\Phi}(\|\sigma_{ik}(\mu)\|)$ auf E nicht abnehmend der Grenze $\boldsymbol{\Phi}(\|\sigma_{ik}(+\infty)\|) = 1$ zustrebt. Wir nehmen also die Definition (312) für alle stetigen Funktionen $\varphi(\mu)$ in Anspruch. Ist $\varphi(\mu)$ reellwertig, so ist $\varphi(\mathfrak{A})$ gewiß Hermitesch. Man beweist leicht unter Benutzung des Weierstraßschen Approximationssatzes, daß (313) auch dann gilt, wenn $\varphi(\mu)$ und $\psi(\mu)$ beliebige stetige Funktionen bezeichnen. [Den Beweis hat man dabei wegen der möglicherweise schwachen Konvergenz (vgl. S. 134) in zwei Schritten zu führen, und zwar unter Benutzung des Majorantenkriteriums: Man zeigt zunächst bei festem φ, daß (313) gilt, wenn ψ lediglich stetig ist, φ jedoch ein Polynom darstellt; man betrachtet sodann bei festem ψ, das jetzt schon beliebig ist, eine gleichmäßig konvergente Folge von Polynomen φ.] — Wegen (312) bzw. (313) gilt

$$(314) \qquad \varphi(\mathfrak{A})\,\psi(\mathfrak{A}) = \varphi\psi(\mathfrak{A}); \qquad c_1\varphi(\mathfrak{A}) + c_2\varphi(\mathfrak{A}) = (c_1\varphi_1 + c_2\varphi_2)\,\mathfrak{A},$$

unter $\varphi\psi(\mathfrak{A}) = \psi\varphi(\mathfrak{A})$ die Matrix verstanden, welche gemäß (312) der Funktion $\psi(\mu)\,\varphi(\mu)$ zugeordnet ist; ganz analog ist die Bedeutung der zweiten Relation (314), wobei die c konstant sind. Der Funktion $\varphi \equiv 1$ ist die Einheitsmatrix, und allgemeiner der n-ten Potenz von μ die n-te Potenz von $\mathfrak{A}$ zugeordnet. — Die Zuordnung ist nach (314) distributiv und „multiplikativ". Sie ist ferner, wie wir gesehen haben, derart, daß $\mathsf{M}(\varphi(\mathfrak{A})) \leqq \mathfrak{m}(\varphi)$ gilt, unter $\mathfrak{m}(\varphi)$ die obere Grenze (Maximum) der Funktion $|\varphi(\mu)|$ verstanden. — Die Zuordnung ist stetig: bezeichnet $\varphi_1, \varphi_2, \ldots$ eine Folge von gegen φ gleichmäßig konvergierenden stetigen Funktionen, so gilt $\varphi_n(\mathfrak{A}) \to \varphi(\mathfrak{A})$, genauer $\mathsf{M}(\varphi_n(\mathfrak{A}) - \varphi(\mathfrak{A})) \to 0$, dies zwar wegen

$$(315) \qquad \mathsf{M}(\varphi_n(\mathfrak{A}) - \varphi(\mathfrak{A})) = \mathsf{M}((\varphi_n - \varphi)\,\mathfrak{A}) \leqq \mathfrak{m}(\varphi_n - \varphi)$$

und da nach Voraussetzung $\mathfrak{m}(\varphi_n - \varphi) \rightarrow 0$. — Übrigens hat F. Riesz diese Zuordnung auf Funktionen φ übertragen, die nicht notwendig stetig, ja nicht einmal halbstetig sind, sondern nur als Summe einer von oben und einer von unten halbstetigen Funktion dargestellt werden können. Bei der „Fortsetzung" von Funktionalen spielt diese Funktionenklasse auch sonst methodisch eine Rolle, die auf dem Umstand beruht, daß diese Funktionen es sind, die durch *absolut* konvergente Polynomreihen (oder, was dasselbe ist, durch absolut konvergente Reihen von willkürlichen stetigen Funktionen), die nicht auch gleichmäßig zu konvergieren brauchen, dargestellt werden können. Da die von rechts stetigen Funktionen von beschränkter Schwankung gewiß dieser Klasse angehören, so konnte F. Riesz eine einfache Herleitung der Formeln (312), (313) usf. auf dieses Zuordnungsprinzip gründen. Er hat auch einen anderen Weg eingeschlagen, der die Frage, mit Rücksicht auf (305), mit dem Momentenproblem in Zusammenhang bringt. Wir werden später bei den unitären Matrizen ähnlich verfahren.

Der multiplikative und distributive Charakter der Zuordnung ergab sich aus der „infinitesimalen Integrabilitätsbedingung" (294). Man kann sich die Frage stellen, inwiefern das Bestehen der „integralen Integrabilitätsbedingung" (314) dasjenige einer „infinitesimalen" bedingt. Diese Frage hängt, wie an anderer Stelle näher ausgeführt werden soll, mit einem Problem betreffend konvexe Funktionale zusammen.

Das Korrespondenzprinzip kann zur Integration des Differentialsystems $\dot{\mathfrak{y}}_t = \mathfrak{A}\mathfrak{y}_t$ verwendet werden. Die allgemeine Lösung ist wieder $\mathfrak{y}_t = e^{\mathfrak{A}t}\mathfrak{y}_0$, d. h.

$$y_i = \sum_{k=1}^{\infty} y_k(0) \int_{-\infty}^{+\infty} e^{\mu t}\, d\sigma_{ik}(\mu),$$

wobei $\sum |y_i(0)|^2 < +\infty$ vorausgesetzt ist. Ebenso ist die allgemeine Lösung des Differentialsystems $\ddot{\mathfrak{x}}_t = \mathfrak{A}\mathfrak{x}_t$ oder $\ddot{x}_i = \sum_{k=1}^{\infty} a_{ik} x_k$ darstellbar durch

$$(316)\quad x_i(t) = \sum_{k=1}^{\infty} \left(x_k(0) \int_{-\infty}^{+\infty} \cos\frac{t}{\sqrt{-\mu}}\, d\sigma_{ik}(\mu) + \dot{x}_k(0) \int_{-\infty}^{+\infty} \sqrt{-\mu}\, \sin\frac{t}{\sqrt{-\mu}}\, d\sigma_{ik}(\mu) \right),$$

wenn von $\mathfrak{x}_0$ und $\dot{\mathfrak{x}}_0$ quadratische Konvergenz angenommen wird. — Daraus ist ersichtlich, daß das gemäß $\ddot{\mathfrak{x}} = \mathfrak{A}\mathfrak{x}$ schwingende System nur solche ebenen „Wellen"[1])

$$(317)\qquad \| d\sigma_{ik}(\mu) \| \cos\frac{t}{\sqrt{-\mu}}, \qquad \| d\sigma_{ik}(\mu) \| \sqrt{-\mu}\, \sin\frac{t}{\sqrt{-\mu}}$$

[1]) Im Spezialfall, wo σ_{ik} nur von der Differenz $i - k$ abhängt, handelt es sich wirklich um Wellen.

emittieren kann, deren μ im Spektrum liegt; denn sonst gilt

$$\int_{\mu-\varepsilon}^{\mu+\varepsilon} \ldots d\sigma_{ik}(\mu) \equiv 0; \qquad i, k = 1, 2, \ldots$$

für hinreichend kleine $\varepsilon (> 0)$. Daher eben der Name Spektrum (Punktspektrum = Linienspektrum, Streckenspektrum = Bandenspektrum. NB. die Formel (316) gilt freilich auch bei endlichen Matrizen $\mathfrak{A}$). Die ebenen Wellen (317), aus welchen sich die allgemeine Lösung (316) gemäß dem Superpositionsprinzip aufbaut, sind dabei — als „Differentiallösungen" — mindestens *formale* Lösungen des Differentialsystems. Vgl. hierzu S. 206 weiter unten.

Die den Mathematikern längst bekannte Korrespondenz zwischen Matrizen und Funktionen enthält, wie wir jetzt sehen werden, das Übertragungsprinzip, welches von der Heisenbergschen Theorie zur Schrödingerschen hinüberführt.

§ 83. Funktionale Gruppenmatrizen.

Wir haben vorher unter Zugrundelegung einer Hermiteschen Matrix einer stetigen Funktion φ gemäß (312), (313) eine Matrix zugeordnet. Die Zuordnung war distributiv, multiplikativ und stetig. Wir wollen jetzt diese Zuordnung in einem wichtigen Spezialfall näher verfolgen. Wir betrachten also wieder eine Abbildung, welche jeder stetigen Funktion φ eine Matrix $\mathfrak{A}(\varphi)$ zuordnet, die beschränkt, für reellwertige φ von Hermiteschem Typus, für $\varphi \equiv 1$ die Einheitsmatrix ist; und zwar soll die Abbildung distributiv und multiplikativ sein,

$$(318) \qquad c_1 \mathfrak{A}(\varphi) + c_2 \mathfrak{A}(\psi) = \mathfrak{A}(c_1 \varphi + c_2 \psi); \qquad \mathfrak{A}(\varphi)\mathfrak{A}(\psi) = \mathfrak{A}(\varphi \psi),$$

und endlich die Stetigkeitseigenschaft haben, daß aus $\mathfrak{m}(\varphi_n - \varphi) \to 0$ immer $\mathsf{M}(\mathfrak{A}(\varphi_n - \varphi)) \to 0$ folgt. — Wir fragen auch diesmal nicht nach der „infinitesimalen" Struktur der *allgemeinsten* Realisierung einer solchen Zuordnung, sondern zeigen nur folgendes: Setzt man

$$(319) \qquad \mathfrak{A}(\varphi) = \left\| \int_{-\infty}^{+\infty} \varphi(\mu)\, \theta_i(\mu)\, \bar{\theta}_k(\mu)\, d\sigma(\mu) \right\|,$$

wobei $\sigma(\mu)$ eine reellwertige, außerhalb eines hinreichend großen Intervalles konstante Funktion von beschränkter Schwankung und $\theta_1, \theta_2, \ldots$ ein stetiges Funktionssystem bedeutet, für welches die Orthogonalitätsbeziehung

$$(320) \qquad \int_{-\infty}^{+\infty} \theta_i(\mu)\, \bar{\theta}_k(\mu)\, d\sigma(\mu) = \delta_{ik}; \qquad i, k = 1, 2, \ldots$$

und für alle stetigen Funktionen f und g die „Vollständigkeitsrelation"

$$(321)\qquad \sum_{j=1}^{\infty}\int_{-\infty}^{+\infty} f(\mu)\bar{\theta}_j(\mu)\,d\sigma(\mu)\cdot\int_{-\infty}^{+\infty} g(\mu)\,\theta_j(\mu)\,d\sigma(\mu)=\int_{-\infty}^{+\infty} f(\mu)\,g(\mu)\,d\sigma(\mu)$$

(die also nicht notwendig im Sinne des Fischer-F. Rieszschen Satzes gemeint ist) gilt, so sind unsere Forderungen gewiß erfüllt. Eine wie (319) gebaute Matrix wollen wir mit Rücksicht auf eine Andeutung von O. Toeplitz, welche die Bedeutung der Zuordnung im Sinne von Frobenius in Evidenz setzt, als eine funktionale Gruppenmatrix bezeichnen. Selbstverständlich handelt es sich dabei nicht um *eine* Matrix, sondern um eine *Klasse* von Matrizen, so daß wir genauer von dem durch $\{\theta_n;\sigma\}$ festgelegten Körper von multiplikativen Klassenmatrizen zu sprechen haben.

Wegen (320) ist $\mathfrak{A}(1)=\mathfrak{E}$, und $\mathfrak{A}(\varphi)$ ist gewiß Hermitesch, wenn φ reellwertig ist; denn $\sigma(\mu)$ ist nach Voraussetzung reellwertig. Der distributive Charakter der Zuordnung ist selbstverständlich. Die zweite unter (318) stehende Beziehung verifiziert sich unter Benutzung von (321) wie folgt:

$$(322)\qquad \mathfrak{A}(\varphi)\,\mathfrak{A}(\psi)=\left\|\int_{-\infty}^{+\infty}\varphi\,\theta_i\bar{\theta}_k\,d\sigma\right\|\,\left\|\int_{-\infty}^{+\infty}\psi\,\theta_i\bar{\theta}_k\,d\sigma\right\|$$

$$=\left\|\sum_{j=1}^{\infty}\int_{-\infty}^{+\infty}\varphi\,\theta_i\bar{\theta}_j\,d\sigma\int_{-\infty}^{+\infty}\psi\,\theta_j\bar{\theta}_k\,d\sigma\right\|=\left\|\int_{-\infty}^{+\infty}\varphi\,\psi\,\theta_i\bar{\theta}_k\,d\sigma\right\|=\mathfrak{A}(\varphi\,\psi)\,.$$

Es gilt endlich auf E

$$(323)\qquad \left|\sum_{i=1}^{n}\sum_{k=1}^{n}x_i\bar{x}_k\int_{-\infty}^{+\infty}\varphi\,\theta_i\bar{\theta}_k\,d\sigma\right|=\left|\int_{-\infty}^{+\infty}\varphi\left|\sum_{i=1}^{n}x_i\theta_i\right|^2d\sigma\right|$$

$$\leqq \mathfrak{m}(\varphi)\int_{-\infty}^{+\infty}\left|\sum_{i=1}^{n}x_i\theta_i\right|^2d\sigma=\mathfrak{m}(\varphi)\sum_{i=1}^{n}\sum_{k=1}^{n}x_i\bar{x}_k\int_{-\infty}^{+\infty}\theta_i\bar{\theta}_k\,d\sigma$$

$$=\mathfrak{m}(\varphi)\sum_{i=1}^{n}|x_i|^2\leqq\mathfrak{m}(\varphi),$$

folglich ist $\mathsf{M}(\mathfrak{A}(\varphi))\leqq\mathfrak{m}(\varphi)$ und daraus folgt wieder die Stetigkeit der Zuordnung ebenso wie bei (315). Die Matrix ist, wir wir soeben gesehen haben, beschränkt, ihre Kopplungsform ist, wie man aus (323) leicht erkennt,

$$(324)\qquad \Phi(\mathfrak{A}(\varphi))=\int_{-\infty}^{+\infty}\varphi(\mu)\left|\sum_{i=1}^{\infty}x_i\theta_i(\mu)\right|^2d\sigma(\mu)\,.$$

Ist λ eine Zahl, die nicht im Wertbereich der Funktion $\varphi(\mu)$ liegt, so ist die Funktion $\frac{1}{\lambda - \varphi(\mu)}$ stetig. Wegen $(\lambda - \varphi)\frac{1}{\lambda - \varphi} = 1$ ist $\mathfrak{A}(\lambda - \varphi)\,\mathfrak{A}\left(\frac{1}{\lambda - \varphi}\right) = \mathfrak{E}$, d. h. $\mathfrak{A}\left(\frac{1}{\lambda - \varphi}\right)$ ist die (offenbar sowohl hintere als auch vordere) Reziproke von $\mathfrak{A}(\lambda - \varphi) = \lambda\,\mathfrak{A}(1) - \mathfrak{A}(\varphi) = \lambda\,\mathfrak{E} - \mathfrak{A}(\varphi)$; m. a. W., die Resolvente von $\mathfrak{A}(\varphi)$ ist

$$\| R_{ik}(\lambda) \| = \mathfrak{A}\left(\frac{1}{\lambda - \varphi}\right) = \left\| \int_{-\infty}^{+\infty} \frac{\theta_i(\mu)\,\bar{\theta}_k(\mu)\,d\sigma(\mu)}{\lambda - \varphi(\mu)} \right\|. \tag{325}$$

Die Matrix $\mathfrak{A}(\varphi)$ ist dabei, wenn φ nicht reellwertig ist, nicht notwendig Hermitesch, jedoch gewiß normal. Denn die begleitende Matrix von $\mathfrak{A}(\varphi)$ ist, wie man sich leicht überzeugt, die Matrix $\mathfrak{A}(\bar{\varphi})$, welche der konjugierten Funktion $\bar{\varphi}(\mu)$ zugeordnet ist, so daß mit Rücksicht auf den multiplikativen Charakter der Zuordnung die beiden Normen von $\mathfrak{A}(\varphi)$ identisch, nämlich gleich $\mathfrak{A}(|\varphi|^2)$ sind. — Ist $\varphi(\mu)$ reellwertig, so wird man, um durch Vergleich mit (293) die Spektralmatrix ablesen zu können, bestrebt sein, in das Integral (325) $\varphi(\mu)$ als unabhängige Variable einzuführen. Es handelt sich also um die Bestimmung der zu $\varphi(\mu)$ inversen Funktion. Beachtet man, daß es stetige Funktionen gibt, die auf keiner Teilstrecke monoton sind, so erkennt man die Ziele, die der Hellingersche Integralbegriff, der u. a. eben zu diesem Zwecke geschaffen wurde, erreichen will und kann. Es kann freilich die „Umkehrung" auch mit Hilfe des Lebesgueschen Integralbegriffes geschehen. — Wir werden später die Verhältnisse an Hand der fastperiodischen Matrizen überblicken können, welche die klassischen (zeitlichen) Schwingungen in die Sprache der unendlichen Matrizen übersetzen.

§ 84. Fortsetzung. Unitäre Äquivalenz der Klassen.

Ist $\mathfrak{A}$ eine funktionale Gruppenmatrix oder eine fastperiodische Matrix, so können die Zerlegungen (278) der Spektralmatrix nach einer Diskussion des Verlaufes der Funktion φ leicht angegeben oder wenigstens beschrieben werden; die einzelnen getrennten Bestandteile in (278) haben dabei für die Darstellung genau die Rolle von irreduziblen Quadraten.

Aus (325) läßt sich übrigens durch eine Verallgemeinerung der in Kap. II durchgeführten Betrachtung entnehmen, daß an der Stelle $\lambda = \lambda_0$ alle Funktionen $R_{ik}(\lambda)$ nur dann regulär sein können, wenn λ_0

nicht dem Wertevorrat der (im Nenner stehenden stetigen) Funktion $\varphi(\mu)$ angehört (vgl. S. 268). Die bei der Herleitung von (325) benutzte Bemerkung kann also umgekehrt werden. Mithin ist das Spektrum der funktionalen Gruppenmatrix $\mathfrak{A}(\varphi)$ mit dem Wertevorrat der Funktion $\varphi(\mu)$ identisch. Mit $\lambda_0 = 0$ folgt hieraus, daß, wenn die Funktion $\varphi(\mu)$ der Null nicht beliebig nahe kommen kann, $\mathfrak{A}(\varphi)$ wegen (325) die beschränkte Reziproke $\mathfrak{A}\left(\frac{1}{\varphi}\right)$ hat, während sonst keine beschränkte Reziproke vorhanden ist (ebenso, wie bei der Funktion φ selbst).

Ist

$$(326) \qquad \mathfrak{B}(\varphi) = \left\| \int_{-\infty}^{+\infty} \varphi(\mu)\, \chi_i(\mu)\, \bar{\chi}_k(\mu)\, d\sigma(\mu) \right\|$$

eine andere Klasse von funktionalen Gruppenmatrizen, wobei die Belegung σ dieselbe wie in (319) ist, und setzt man

$$(327) \qquad \mathfrak{U} = \left\| \int_{-\infty}^{+\infty} \theta_i(\mu)\, \bar{\chi}_k(\mu)\, d\sigma(\mu) \right\|,$$

so folgt, genau so wie auf S. 66, daß

$$(328) \qquad \mathfrak{U}\,\mathfrak{A}(\varphi)\,\mathfrak{U}^* = \mathfrak{B}(\varphi), \qquad \mathfrak{U}^*\,\mathfrak{B}(\varphi)\,\mathfrak{U} = \mathfrak{A}(\varphi)$$

gilt. Setzt man $\varphi \equiv 1$, so geht (328) in die definitorische Bedingung des unitären Charakters über, so daß die beiden Matrizen $\mathfrak{A}(\varphi)$, $\mathfrak{B}(\varphi)$ unitär äquivalent sind. Im Rahmen des vorigen Paragraphen entspricht den Schlußbemerkungen von § 7 die leicht beweisbare Identität

$$(329) \qquad \mathfrak{U}\,\varphi(\mathfrak{A})\,\mathfrak{U}^* = \varphi(\mathfrak{U}\,\mathfrak{A}\,\mathfrak{U}^*), \qquad \mathfrak{U}^* = \mathfrak{U}^{-1}.$$

Das Wesentliche bei den multiplikativen Gruppenmatrizen ist der folgende Umstand. Weiß man nur, daß $a_{ik} = \int_{-\infty}^{+\infty} \mu\, d\sigma_{ik}(\mu)$ gilt, so folgt daraus noch nicht, daß $\|\sigma_{ik}\|$ die Spektralmatrix ist. Denn die σ_{ik} sind gewissermaßen unabhängig voneinander. Nimmt man aber an, daß $d\sigma_{ik} = \theta_i \bar{\theta}_k\, d\sigma$ gilt, so hat man nicht mehr ∞^2, sondern nur ∞ Bestimmungsstücke zur Verfügung. Nimmt man außerdem an, daß (320) und (321) gelten, wobei (305) notwendig ist, denn es ist gewiß, da $\left\|\int_{-\infty}^{+\infty} d\sigma_{ik}\right\| = \mathfrak{A}^0 = \mathfrak{E}$ sein muß, so folgt aus (321), daß (305) für $n > 1$, und aus (320), daß (305) auch für $n = 0$ gilt. Kennt man aber alle $\mathfrak{A}^n$; $n = 0, 1, \ldots$, so kann man die C. Neumannsche Reihe bilden,

welche, mit Rücksicht auf die Stieltjessche Umkehrformel, die Spektralmatrix bereits festlegt. Diese Bemerkung rührt von Hilbert selbst her, der die Fourierschen Matrizen eingeführt hat. Hilbert hat dabei vermutet, daß die Fourierschen Matrizen gewissermaßen Normalbestandteile beliebiger Hermiteschen Matrizen sind. Im Anschluß daran ist dann die Hellingersche Zerlegungstheorie entstanden, welche die Spektralmatrix beliebiger Hermiteschen Matrizen aus den Spektralmatrizen getrennter Fouriermatrizen aufzubauen gestattet. Wir kommen darauf noch zurück.

§ 85. Fortsetzung. Polynomische Klassen.

Wegen (328) sind die verschiedenen Klassen funktionaler Gruppenmatrizen gleichwertig, so daß es zweckmäßig ist, die Betrachtungen von § 30 auf den hier vorliegenden Fall zu übertragen. Es sei $\sigma(\mu) \geqq 0$ eine nicht abnehmende Funktion, die keine Treppenfunktion ist, also unendlich viele Wachstumstellen hat, die jedoch alle ganz im Endlichen liegen mögen, so daß σ außerhalb eines hinreichend großen Intervalles konstant ist. Dann gibt es Polynome $\theta_1, \theta_2, \ldots,$ wobei θ_i vom genauen Grade $i-1$ ist, derart, daß die Bedingungen (320), (321) in bezug auf die vorgegebene Belegung σ erfüllt sind.

Zunächst ist $\int\limits_{-\infty}^{+\infty} \Theta(\mu)\, d\sigma(\mu) \neq 0$, wenn $\Theta(\mu) \not\equiv 0$ ein Polynom bedeutet, das für alle $\mu \gtreqless 0$ durchweg $\geqq 0$ ist. Da nämlich $\Theta(\mu) \geqq 0$ ein Polynom ist und $\sigma(\mu)$ nicht nur endlich viele Wachstumsstellen hat, so gibt es ein $\delta > 0$ und dazu ein Intervall $[\alpha, \beta]$ derart, daß für $\alpha \leqq \mu \leqq \beta$ das Polynom $\Theta(\mu) \geqq \delta$, und zugleich $\int\limits_{\alpha}^{\beta} d\sigma(\mu) > 0$, also $\int\limits_{-\infty}^{+\infty} \Theta\, d\sigma \geqq \int\limits_{\alpha}^{\beta} \Theta\, d\sigma \geqq \delta \int\limits_{\alpha}^{\beta} d\sigma > 0$ ausfällt. Man erkennt hierauf nach dem Gram-Schmidtschen Orthogonalisierungsverfahren durch vollständige Induktion (vgl. S. 12), daß Polynome $\theta_1, \theta_2, \ldots,$ wobei θ_i genau vom Grade $i-1$ und $\theta_1 \not\equiv 0$ ist, sukzessive derart bestimmt werden können, daß (320) gilt. Es folgt hieraus wie auf S. 67, daß jedes Polynom $\Theta(\mu)$ vom Grade $n-1$ in der Gestalt $c_1\theta_1 + \cdots + c_n\theta_n$ dargestellt werden kann, und zwar nur auf eine Art und Weise, indem $c_\nu = \int\limits_{-\infty}^{+\infty} \Theta\, \bar{\theta}_\nu\, d\sigma$ sein muß; und daß ferner $\int\limits_{-\infty}^{+\infty} \Theta\, \bar{\theta}_\nu\, d\sigma = 0$ gilt, wenn $\nu \geqq n$ ist. Mithin ist (321) für den Fall, daß φ und ψ Polynome sind, eine leicht erkennbare Identität, und daraus folgt unter Benutzung der

beim Beweise von (313) angedeuteten Schlußweise, daß (321) für jedes Paar φ, ψ von stetigen Funktionen gilt. Denn nach Weierstraß kann jede stetige Funktion auf jedem endlichen Intervalle durch Polynome gleichmäßig approximiert werden, und hier handelt es sich nur um ein endliches Intervall, da σ sonst konstant ist.

Wir wollen fortan mit $\sigma \geqq 0$ eine Funktion bezeichnen, die nicht abnehmend, von rechts stetig und außerhalb eines hinreichend großen Intervalles konstant ist; $\{\theta_i\}$, $\{\chi_i\}$ bezeichnen zugehörige, den Bedingungen (320), (321) [bzw. den beiden analogen, dem System $\{\chi_i\}$ entsprechenden Bedingungen] genügende Funktionensysteme, wobei θ_i ein Polynom vom genauen Grade $i-1$ ist, so daß (siehe S. 68)

$$\int_{-\infty}^{+\infty} \Theta(\mu)\,\theta_n(\mu)\,d\sigma(\mu)$$

gewiß verschwindet oder aber nicht verschwindet, je nachdem der genaue Grad des Polynoms $\Theta(\mu)$ kleiner als n oder $= n$ ist. Insbesondere ist also

$$\int_{-\infty}^{+\infty} \mu\,\theta_i(\mu)\,\bar{\theta}_k(\mu)\,d\sigma(\mu)$$

gewiß $= 0$ oder aber $\neq 0$, je nachdem $1+i > k$ oder $1+k > i$, bzw. $1+i = k$ oder $1+k = i$ gilt. Denn $\{\bar{\theta}_i\}$ genügt denselben Bedingungen wie $\{\theta_i\}$.

§ 86. Fortsetzung. Fouriersche Matrizen.

Die beschränkten Hermiteschen Matrizen $\mathfrak{A}(\varphi)$, $\mathfrak{B}(\varphi)$ unter (319), (326), die zu $\varphi \equiv \mu$ gehören,

$$(330)\quad \mathfrak{A}(\mu) = \int_{-\infty}^{+\infty} \mu\,\chi_i(\mu)\,\bar{\chi}_k(\mu)\,d\sigma(\mu), \qquad \mathfrak{B}(\mu) = \int_{-\infty}^{+\infty} \mu\,\theta_i(\mu)\,\bar{\theta}_k(\mu)\,d\sigma(\mu),$$

nennen wir wieder Fouriersche Matrizen. Als Jacobische Matrizen bezeichnen wir hingegen diejenigen beschränkten Hermiteschen Matrizen, für welche

$$(331)\quad b_{ik} = 0 \quad \text{für} \quad |i-k| > 1 \quad \text{und} \quad b_{i\,i-1} = \bar{b}_{i-1\,i} \neq 0; \quad b_{ii} \gtreqless 0$$

ausfällt. Eine polynomische Fouriersche Matrix $\mathfrak{B}(\mu)$ ist also gewiß eine Jacobische Matrix. Wir werden sogleich sehen, daß auch die Umkehrung gilt. — Jede Fouriersche Matrix $\mathfrak{A}(\mu)$ kann in eine polynomische Fouriersche Matrix unitär transformiert werden. Denn einer-

seits kann man zur Belegung σ von $\mathfrak{A}$ eine den Bedingungen (320), (321) genügende Polynomfolge $\{\theta_i\}$ bestimmen, wobei θ_i genau vom Grade $i-1$ ist, und andererseits gilt (328). Es sei jedoch betont, daß die Übereinstimmung der Belegungen zweier Fourierschen Matrizen zwar eine hinreichende, jedoch keine notwendige Bedingung für ihre unitäre Äquivalenz darstellt. Die Verhältnisse sind nämlich die gleichen wie bei den endlichen Matrizen (S. 63ff.).

Die Spektralmatrix der Fourierschen Matrizen kann man ohne weiteres angeben. Die Resolvente von $\mathfrak{A}(\mu)$ ist nämlich wegen (325)

$$\| R_{ik}(\lambda) \| = \mathfrak{A}\left(\frac{1}{\lambda-\mu}\right) = \left\| \int_{-\infty}^{+\infty} \frac{\chi_i(\mu)\,\bar{\chi}_k(\mu)\,d\sigma(\mu)}{\lambda-\mu} \right\|, \tag{332}$$

wofür, wenn

$$\sigma_{ik}(\mu) = \int_{-\infty}^{\mu} \chi_i(\mu)\,\bar{\chi}_k(\mu)\,d\sigma(\mu) \tag{333}$$

gesetzt wird, nach § 40 auch

$$R_{ik}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\sigma_{ik}(\mu)}{\lambda-\mu} \tag{293$'$}$$

geschrieben werden kann. Wegen (333) ist $\sigma_{ik}(-\infty) = 0$, und $\sigma_{ik}(\mu)$ ist durchweg von rechts stetig, da $\sigma(\mu)$ nach Voraussetzung von rechts stetig ist. Wegen (293) und mit Rücksicht auf die Stieltjessche Umkehrformel muß also die durch (333) gegebene Matrix $\| \sigma_{ik}(\mu) \|$ die Spektralmatrix sein. Denn es gibt nur eine Resolvente und nur eine Spektralmatrix, und für die Spektralmatrix ist (293) wegen (293)′ gewiß richtig.

§ 87. Fortsetzung. Jacobische Matrizen.

Wir zeigen jetzt, daß jede Jacobische Matrix eine zu Polynomen gehörige Fouriersche Matrix ist. Ist nämlich $\| a_{ik} \|$ eine Jacobische Matrix, so ist es auch $\| a_{ik} \|_{[n]}$, also ist die Resolvente $\| R_{ik}^{(n)}(\lambda) \|_{(n)}$ dieses Abschnittes, wie wir wissen (§ 32), von der Struktur

$$R_{ik}^{(n)}(\lambda) = \int_{-\infty}^{+\infty} \frac{\chi_i(\mu)\,\bar{\chi}_k(\mu)\,d\varrho^{(n)}(\mu)}{\lambda-\mu} = \int_{-\infty}^{+\infty} \frac{d\sigma_{ik}^{(n)}(\mu)}{\lambda-\mu}; \tag{334}$$

$$\sigma_{ik}^{(n)}(\mu) = \int_{-\infty}^{\mu} \chi_i(\mu)\,\bar{\chi}_k(\mu)\,d\varrho^{(n)}(\mu); \qquad i,k = 1, 2, \ldots, n, \tag{335}$$

wobei $\varrho^{(n)}(\mu)$ eine nicht abnehmende Treppenfunktion mit genau n Sprungstellen und $\chi_i(\mu)$ ein Polynom genau vom Grade $i-1$ bezeichnet; es gilt

$$(336)\qquad \int_{-\infty}^{+\infty} \chi_i(\mu)\,\bar{\chi}_k(\mu)\,d\varrho^{(n)}(\mu) = \delta_{ik}; \qquad i, k = 1, 2, \ldots, n,$$

und χ_i hängt mit Rücksicht auf (115) von n nicht ab. Es muß also nach S. 81 eine Funktion $\sigma_{ik}(\mu)$ von beschränkter Schwankung geben derart, daß an allen Stetigkeitsstellen von $\sigma_{ik}(\mu)$

$$(337)\qquad \int_{-\infty}^{\mu} \chi_i(\mu)\,\bar{\chi}_k(\mu)\,d\varrho^{(n)}(\mu) \to \sigma_{ik}(\mu)$$

gilt, und zwar ist $\|\sigma_{ik}(\mu)\|$ die Spektralmatrix der unendlichen Matrix, sofern die $\sigma_{ik}(\mu)$ nötigenfalls von rechts stetig gemacht werden. Da χ_1 eine Konstante $C \neq 0$ ist, so muß insbesondere

$$(338)\qquad \int_{-\infty}^{\mu} d\varrho^{(n)}(\mu) \to \frac{\sigma_{11}(\mu)}{|C|^2}$$

an allen Stetigkeitsstellen von σ_{11} gelten. Wir wählen eine Zahl μ' derart, daß μ' nicht eine Sprungstelle von σ_{11} oder von einem $\sigma_{11}^{(n)}$ ist; diese μ' liegen überall dicht, und sind, in Anbetracht dessen, daß $\varrho^{(n)} \geqq 0$ nicht abnimmt und im negativen Unendlichen verschwindet, gewiß derart, daß

$$(339)\qquad \int_{-\infty}^{\mu'} d\varrho^{(n)}(\mu) = \varrho^{(n)}(\mu'), \qquad \text{also} \qquad \varrho^{(n)}(\mu') \to \sigma(\mu')$$

gilt, wenn $\sigma(\mu) = \frac{\sigma_{11}(\mu)}{|C|^2}$ gesetzt wird. Nach dem Hellyschen Fortpflanzungssatz (S. 78) gilt also $\varrho^{(n)} [\to] \sigma$; aus (337), (339) folgt daher nach dem Hellyschen Satz über gliedweise Integration (S. 89), daß

$$(340)\qquad \int_{-\infty}^{\mu} \chi_i(\mu)\,\bar{\chi}_k(\mu)\,d\sigma(\mu) = \sigma_{ik}(\mu)$$

an allen Stetigkeitsstellen von $\sigma(\mu)$ gilt. Aus demselben Grunde geht (337) in (333) über. Da $\|\sigma_{ik}\|$ die Spektralmatrix ist, so gilt wegen (305), (340) und mit Rücksicht auf § 40

$$a_{ik} = \int_{-\infty}^{+\infty} \mu\,d\sigma_{ik}(\mu) = \int_{-\infty}^{+\infty} \mu\,\chi_i(\mu)\,\bar{\chi}_k(\mu)\,d\sigma(\mu).$$

Hierbei ist $\sigma(\mu)$ nicht abnehmend, kann als von rechts stetig angenommen werden und hat, wie aus der für *alle* i und k gültigen normierten Orthogonalitätsrelation (320), woselbst χ für θ zu setzen ist, leicht entnommen werden kann, unendlich viele Wachstumsstellen, ist also keine Treppenfunktion. Es gilt dabei (321), wie wir wissen (S. 85), von selbst. Mithin ist $\| a_{ik} \|$, wie behauptet, eine zu einem „vollständigen" Polynomsystem gehörige Fouriersche Matrix.

§ 88. Fortsetzung. Laurentsche Matrizen.

Durch einen vorsichtigen, jedoch ganz elementaren Grenzübergang kann daraus, daß (313) und (321) für alle stetigen Funktionenpaare gelten, ohne Mühe geschlossen werden, daß (313) und (321) auch für alle etwa abteilungsweise stetigen, beschränkten Funktionenpaare gelten, deren Unstetigkeitsstellen nicht in eine Sprungstelle von σ fallen. Diese Bedingung reicht hin, um die Existenz des Stieltjesschen Grenzwertes $\lim S$ von S. 83 zu sichern. — Auf diese geläufigen ε-Betrachtungen können wir hier nicht näher eingehen.

Nehmen wir die Zuordnung durch

(341) $$\mathfrak{A}(\varphi) = \left\| \int_0^1 \varphi(\mu)\, \chi_i(\mu)\, \overline{\chi}_k(\mu)\, d\mu \right\|$$

vor, wobei

(342) $$\int_0^1 \chi_i(\mu)\, \overline{\chi}_k(\mu)\, d\mu = \delta_{ik};$$

$$\sum_{j=1}^{\infty} \int_0^1 \varphi(\mu)\, \chi_j(\mu)\, d\mu \cdot \int_0^1 \psi(\mu)\, \overline{\chi}_j(\mu)\, d\mu = \int_0^1 \varphi(\mu)\, \psi(\mu)\, d\mu$$

ist, so daß $\sigma(\mu)$ durchweg stetig, für $\mu \geqq 1$ und für $\mu \leqq 0$ konstant, für $0 \leqq \mu \leqq 1$ aber $\equiv \mu$ ist, so können wir freilich beliebige beschränkte und integrable Funktionen $\varphi(\mu)$ zulassen. Die Formel (342) gilt für alle im Riemannschen Sinne integrierbaren beschränkten Funktionen φ, ψ, wenn sie für alle stetigen Funktionen gilt. Man kann aber auch fordern, daß $\{\chi_i\}$ im Sinne des Riesz-Fischerschen Satzes vollständig ist; dann gilt (342) für alle Paare nebst ihren Quadraten im Lebesgueschen Sinne integrierbarer Funktionen. Der distributive und multiplikative Charakter der Zuordnung bleibt dabei bestehen und es gilt auch jetzt

(343) $$\mathsf{M}(\mathfrak{A}(\varphi)) \leqq \mathfrak{m}(\varphi),$$

unter $\mathfrak{m}(\varphi)$ die obere Grenze von $|\varphi(\mu)|$ für $0 \leqq \mu \leqq 1$ verstanden. Wir wollen unter $\mathfrak{m}(\varphi)$ die „eigentliche" obere Grenze, nämlich die kleinste Zahl ($\leqq +\infty$) verstehen, für welche $|\varphi(\mu)| \leqq \mathfrak{m}(\varphi)$ abgesehen von einer μ-Nullmenge richtig ist. Es läßt sich dann zeigen, daß $\mathsf{M}(\mathfrak{A}(\varphi))$ nicht nur $\leqq \mathfrak{m}(\varphi)$, sondern genau $= \mathfrak{m}(\varphi)$ ist. Insbesondere ist also die Matrix dann und nur dann beschränkt, wenn die Funktion „eigentlich" beschränkt ist.

Toeplitz hat die Theorie insbesondere derjenigen Matrizen ausgebildet, die zu gewöhnlichen (trigonometrischen) Fourierreihen gehören. Die Numerierung läuft dabei von $-\infty$ bis $+\infty$, doch kann man hiervon zu der Bezifferung von § 83 übergehen. Den Laurentschen Matrizen entsprechen spektral die endlichen „Zyklanten" in einem analogen Sinne, wie die unendlichen Jacobischen Matrizen die Grenzfälle der endlichen sind (S. 191). Man erkennt dies, indem man die Integraldarstellung der Fourierkoeffizienten durch Nährungswerte ersetzt, die zu einer äquidistanten trigonometrischen Interpolation gehören.

Wir werden sehen, daß diese Theorie, wenigstens in ihren Hauptzügen, auf fastperiodische Funktionen ausgedehnt werden kann (soweit ich aus einer Andeutung schließen kann, sind Dirichletsche Reihen in diesem Sinne bereits von Toeplitz selbst betrachtet worden).

§ 89. Unitäre Kovarianz der Spektralmatrix.

Nach diesem Exkurs nehmen wir die Behandlung der allgemeinen Theorie wieder auf. Es sei $\mathfrak{A} = \| a_{iu} \|$ eine beschränkte Hermitesche Matrix, $\mathfrak{R}_\lambda = \| R_{ik}(\lambda) \|$ ihre Resolvente, $\| \sigma_{ik}(\mu) \|$ ihre Spektralmatrix und $\mathfrak{U}$ eine unitäre Matrix. Wegen

$$(\lambda \mathfrak{E} - \mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1})\,\mathfrak{U}\mathfrak{R}_\lambda\mathfrak{U}^{-1} = \mathfrak{U}(\lambda\mathfrak{E} - \mathfrak{A})\,\mathfrak{U}^{-1}\mathfrak{U}\mathfrak{R}_\lambda\mathfrak{U}^{-1}$$
$$= \mathfrak{U}(\lambda\mathfrak{E} - \mathfrak{A})\,\mathfrak{R}_\lambda\mathfrak{U}^{-1} = \mathfrak{U}\mathfrak{E}\mathfrak{U}^{-1} = \mathfrak{U}\mathfrak{U}^{-1} = \mathfrak{E} = \mathfrak{U}^{-1}\mathfrak{U} = \cdots$$

ist $\mathfrak{U}\mathfrak{R}_\lambda\mathfrak{U}^{-1}$ die Resolvente von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$. Wegen $\mathfrak{U}^{-1} = \mathfrak{U}^*$ und mit Rücksicht auf (293) ist also das (i, k)-te Element der Resolvente von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ gleich

$$(344)\qquad \sum_{j=1}^{\infty}\sum_{l=1}^{\infty} u_{ij}\int_{-\infty}^{+\infty}\frac{d\sigma_{jl}(\mu)}{\lambda-\mu}\,\bar{u}_{kl} = \int_{-\infty}^{+\infty}\frac{d\sum_{j=1}^{\infty}\sum_{l=1}^{\infty} u_{ij}\,\sigma_{jl}(\mu)\,\bar{u}_{kl}}{\lambda-\mu} = \int_{-\infty}^{+\infty}\frac{d\sigma_{ik}^*(\mu)}{\lambda-\mu},$$

wenn $\| \sigma_{ik}^*(\mu) \| = \mathfrak{U}\,\| \sigma_{ik}(\mu) \|\,\mathfrak{U}^{-1}$ gesetzt wird; vgl. (268). Die gliedweise Integration rechtfertigt sich dabei wie folgt. Die in (344) linkerhand stehende Doppelsumme ist nach S. 175 gleich

$$(345)\qquad \sum_{j=1}^{\infty} u_{ij}\left(\sum_{l=1}^{\infty}\int_{-\infty}^{+\infty}\frac{d\sigma_{jl}(\mu)}{\lambda-\mu}\,\bar{u}_{kl}\right).$$

Nach S. 164 sind aber die Funktionen $\sum_{l=1}^{n} \sigma_{jl}(\mu)\,\bar{u}_{kl}$; $n = 1, 2, \ldots$ von gleichmäßig beschränkter Schwankung und nach der Schwarzschen Ungleichheit $\leqq 1$, also gleichmäßig beschränkt. Sie streben für $n \to \infty$ der Grenze $\sum_{l=1}^{\infty} \sigma_{jl}(\mu)\,\bar{u}_{kl}$ zu, so daß nach dem Hellyschen Satze über gliedweise Integration die innere Summe in (345) gleich

$$\int_{-\infty}^{+\infty} \frac{d \sum_{l=1}^{\infty} \sigma_{jl}(\mu)\,\bar{u}_{kl}}{\lambda - \mu}$$

ist. Die wiederholte Anwendung derselben Schlußweise ergibt nunmehr die Formel (344), welche, angesichts der Stieltjesschen Umkehrformel und mit Rücksicht auf (293), besagt, daß mit der Matrix $\mathfrak{A}$ sich nicht nur ihre Resolvente, sondern auch ihre Spektralmatrix kogredient transformiert. Nach § 74 und laut der Definition auf S. 177 sind also das Spektrum, das Streckenspektrum, das Punktspektrum und das Häufungsspektrum von $\mathfrak{A}$ unitär invariant. Das Spektrum ist aber nach S. 178 die Gesamtheit der λ-Werte, bei welchen für $\lambda\mathfrak{E} - \mathfrak{A}$ nicht der Fall **1**, mithin, da Fall **3** hier ausgeschlossen ist (S. 139), der Fall **2** von § 60 vorliegt. — Es entsteht nun das Problem, die drei unitär invarianten Bestandteile des Spektrums (für welche also ohne Ausnahme der Fall **2** vorliegt) ebenfalls lösungstheoretisch zu charakterisieren (vgl. § 62).

§ 90. Das Feld der normierten Eigenlösungen.

Wir werden auf S. 198 ff. den Hilbertschen Satz beweisen, demzufolge λ_0 dann und nur dann dem Punktspektrum angehört, wenn die zu $\lambda_0\mathfrak{E} - \mathfrak{A}$ gehörigen homogenen Gleichungen eine vom Nullvektor verschiedene, jedoch quadratisch konvergente Lösung, oder, was dasselbe ist, eine normierte Lösung zulassen.

Diese Eigenschaft erweist sich auch ohne Heranziehung der Spektraltheorie als unitär invariant. Aus $\mathfrak{B}\mathfrak{c} = \mathfrak{o}$ folgt nämlich $\mathfrak{U}(\mathfrak{B}\mathfrak{c}) = \mathfrak{U}\mathfrak{o} = \mathfrak{o}$, also offenbar auch $(\mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1}\mathfrak{U})\mathfrak{c} = \mathfrak{o}$, d. h. $(\mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1})\mathfrak{b} = \mathfrak{o}$, $\mathfrak{b} = \mathfrak{U}\mathfrak{c}$, wobei, da $\mathfrak{U}$ beschränkt ist, der quadratisch konvergente Vektor $\mathfrak{c}$ durch die Transformation mittels $\mathfrak{U}$ mit Rücksicht auf die Schlußbemerkung von § 54 in einen quadratisch konvergenten übergeht, der nur dann $= \mathfrak{o}$ ist, wenn $\mathfrak{c} = \mathfrak{o} =$ Nullvektor war. Aus $\mathfrak{o} = \mathfrak{U}\mathfrak{c}$ folgt in der Tat $\mathfrak{U}^{-1}\mathfrak{o} = \mathfrak{U}^{-1}(\mathfrak{U}\mathfrak{c})$, d. h. $\mathfrak{o} = \mathfrak{E}\mathfrak{c} = \mathfrak{c}$. Setzt man ferner $\mathfrak{B} = \lambda_0\mathfrak{E} - \mathfrak{A}$, so ist

$$\mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1} = \lambda_0\mathfrak{E} - \mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}.$$

Wir können daher, auch *bevor* wir zum Beweis des Hilbertschen Satzes übergehen, die Matrix $\mathfrak{A}$ vereinfachenden unitären Transformationen unterwerfen. Analog ersieht man, daß die Anzahl der zu $\mathfrak{B}\mathfrak{x} = \mathfrak{o}$ gehörigen linear unabhängigen quadratisch konvergenten Eigenlösungen, die $= 0$, endlich und positiv, oder $+\infty$ sein kann, für $\mathfrak{B}$ und $\mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1}$ dieselbe ist. Hierzu ist wieder (vgl. S. 8) folgendes zu bemerken. Es sei Ω ein den Nullvektor enthaltendes „Feld" von quadratisch konvergenten Vektoren $\mathfrak{x}$, so daß mit $\mathfrak{x}_1, \ldots, \mathfrak{x}_m$ immer auch die (der Schwarzschen Ungleichheit zufolge gewiß quadratisch konvergente) endliche homogen-lineare Verbindung $\alpha_1 \mathfrak{x}_1 + \cdots + \alpha_m \mathfrak{x}_m$ der Gesamtheit Ω angehört. Die Gesamtheit der zu $\mathfrak{B}\mathfrak{x} = \mathfrak{o}$ gehörigen quadratisch konvergenten Lösungen hat gewiß diese Eigenschaft. Wir ordnen Ω eine nichtnegative ganze Zahl ν oder $\nu = +\infty$ als den *Rang* von Ω auf die folgende Weise zu. Wir setzen ν dann und nur dann $= 0$, wenn Ω außer dem Nullvektor keinen weiteren Vektor enthält. Ist $\nu > 0$, so sind zwei Fälle möglich. I. Entweder gibt es in Ω eine Anzahl $N\,(> 0)$ von Vektoren $\mathfrak{x}_1^0, \mathfrak{x}_2^0, \ldots, \mathfrak{x}_N^0$, aus welchen jeder andere Vektor des Feldes homogen-linear zusammengesetzt werden kann, während aus $\alpha_1^0 \mathfrak{x}_1^0 + \cdots + \alpha_N^0 \mathfrak{x}_N^0 = \mathfrak{o}$ immer $\alpha_1^0 = \cdots = \alpha_N^0 = 0$ folgt, so daß die N Grundvektoren linear unabhängig sind. Man überzeugt sich leicht, daß N von der Wahl der Grundvektoren nicht abhängt, also durch Ω eindeutig festgelegt wird. Wir setzen den Rang ν gleich N, der also jetzt > 0, aber $< +\infty$ ist. — II. Oder aber es gibt keine solche Anzahl N, sie ist $= +\infty$. Wir setzen dann $\nu = +\infty$. (Es ist dabei zu beachten, daß aus $\nu = +\infty$ noch keinesfalls folgt, daß Ω *alle* quadratisch konvergenten Vektoren enthält, so daß die Bezeichnung Rang bei Hilbertschen Vektorkörpern nur cum grano salis zu verstehen ist; vgl. hingegen S. 10.)

Den Rang von $\lambda\mathfrak{E} - \mathfrak{A}$ wollen wir durch $\nu(\lambda)$ bezeichnen und Vielfachheit von λ nennen. Für $\mathfrak{A}$ und $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ hat λ, wie erwähnt, dieselbe Vielfachheit. — Den erwähnten Hilbertschen Satz können wir offenbar so aussprechen: der Punkt λ gehört dann und nur dann dem Punktspektrum an, wenn $1 \leqq \nu(\lambda) \leqq +\infty$ gilt. Gehört also λ dem Häufungsspektrum oder dem Streckenspektrum oder beiden an, ohne zugleich dem Punktspektrum anzugehören (vgl. S. 161), so ist $\nu(\lambda) = 0$, und es liegt dabei für $\lambda\mathfrak{E} - \mathfrak{A}$, ebenso wie im Falle $\nu(\lambda) \geqq 1$, der Fall **2** von S. 138 vor. Gehört λ dem Spektrum überhaupt nicht an, so ist freilich $\nu(\lambda) = 0$ ebenfalls richtig (vgl. S. 141), es gibt jedoch dann eine und nur eine beschränkte Reziproke, es liegt also der Fall **1** vor.

§ 91. Hauptachsentransformation des punktspektralen Bestandteiles.

Es bezeichne wieder $\|\tilde{\sigma}_{ik}(\mu)\|$ den stetigen, $\|\hat{\sigma}_{ik}(\mu)\|$ den sprunghaften Bestandteil von $\|\sigma_{ik}(\mu)\|$, d. h. der Spektralmatrix von $\mathfrak{A}$. Dann ist, wie wir wissen (S. 165 bzw. 194), $\|\tilde{\sigma}^*_{ik}(\mu)\| = \mathfrak{U}\|\tilde{\sigma}_{ik}(\mu)\|\mathfrak{U}^{-1}$ der stetige, $\|\hat{\sigma}_{ik}(\mu)\| = \mathfrak{U}\|\hat{\sigma}_{ik}(\mu)\|\mathfrak{U}^{-1}$ der sprunghafte Teil von

$$(346) \qquad \|\sigma^*_{ik}(\mu)\| = \mathfrak{U}\|\sigma_{ik}(\mu)\|\mathfrak{U}^{-1},$$

d. h. der Spektralmatrix von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$. Es bezeichne $\lambda_1, \lambda_2, \ldots$ das Punktspektrum von $\mathfrak{A}$, also auch von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$. Unter Umständen gibt es überhaupt kein Punktspektrum oder auch überhaupt kein Streckenspektrum. Wir wollen vom allgemeinen Falle sprechen und, um Vorstellungen zu fixieren, annehmen, daß es sogar unendlich viele λ_j gibt. Sie sind dadurch charakterisiert, daß sie die Sprungstellen von mindestens einem $\sigma_{ik}(\mu)$ oder $\hat{\sigma}_{ik}(\mu)$; $i, k = 1, 2, \ldots$ darstellen, und diese Eigenschaft ist unitär invariant. Die Zahlen λ_j sind nach Definition voneinander durchweg verschieden, ihr „Gewicht" soll eben jetzt, an der Hand des Hauptachsenproblems erklärt werden. Man hat für $\mathfrak{A}$

$$(347) \qquad \sigma_{ik}(\mu) = \tilde{\sigma}_{ik}(\mu) + \hat{\sigma}_{ik}(\mu), \qquad \hat{\sigma}_{ik}(\mu) = \sum_{\lambda_j \leqq \mu} \underset{j}{\Delta}\sigma_{ik},$$

und für $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ entsprechend

$$(348) \qquad \sigma^*_{ik}(\mu) = \tilde{\sigma}^*_{ik}(\mu) + \hat{\sigma}^*_{ik}(\mu), \qquad \hat{\sigma}^*_{ik}(\mu) = \sum_{\lambda_j \leqq \mu} \underset{j}{\Delta}\sigma^*_{ik},$$

wobei $\underset{j}{\Delta} = \underset{\lambda_j}{\Delta}$ gesetzt ist.

Wir nehmen nun auf den Satz von S. 163 Bezug: wir können die unitäre Matrix $\mathfrak{U}$ unabhängig von μ derart wählen, daß die Zerlegung (269) der Spektralmatrix $\|\sigma^*_{ik}(\mu)\|$ von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ folgendermaßen beschaffen ist: In der Zerlegung

$$(348)' \qquad \|\sigma^*_{ik}(\mu)\| = \|\tilde{\sigma}^*_{ik}(\mu)\| + \Big\|\sum_{\lambda_j \leqq \mu} \underset{\lambda_j}{\Delta}\sigma^*_{ik}\Big\|$$

sind die Matrizen $\|\tilde{\sigma}^*_{ik}(\mu)\|$, $\|\underset{\lambda_1}{\Delta}\sigma^*_{ik}\|$, $\|\underset{\lambda_2}{\Delta}\sigma^*_{ik}\|, \ldots$ paarweise getrennt für alle μ und es ist $\underset{j}{\Delta}\sigma^*_{ik} = \gamma_i^{(j)}\delta_{ik}$, wobei $\gamma_i^{(j)}$ entweder $= 0$ oder $= 1$ ist. Bezeichnet daher $\|R^*_{ik}(\lambda)\|$ die Resolvente von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ — eine Verwechslung mit der begleitenden Matrix $\|R_{ik}(\lambda)\|^*$ der Resolvente von $\mathfrak{A}$ ist wohl nicht zu befürchten —, so ist wegen (293)

$$(359) \qquad \|R^*_{ik}(\lambda)\| = \|\tilde{R}^*_{ik}(\lambda)\| + \sum_{j=1}^{\infty}\left\|\frac{\gamma_i^{(j)}\delta_{ik}}{\lambda - \lambda_j}\right\|,$$

wobei die Matrizen rechterhand für alle λ paarweise getrennt sind und

(360) $$\| \tilde{R}_{ik}^{*}(\lambda) \| = \left\| \int_{-\infty}^{+\infty} \frac{d\tilde{\sigma}_{ik}^{*}(\mu)}{\lambda - \mu} \right\|$$

gesetzt ist; vgl. nämlich S. 175. Aus (305) folgt mit $\nu = 1$ ebenso

(361) $$\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1} = \| \int_{-\infty}^{+\infty} \mu \, d\tilde{\sigma}_{ik}^{*}(\mu) \| + \sum_{j=1}^{\infty} \| \gamma_i^{(j)} \lambda_j \delta_{ik} \|,$$

und mit $\nu = 0$ wegen $\mathfrak{U}\mathfrak{A}^0\mathfrak{U}^{-1} = \mathfrak{U}\mathfrak{E}\mathfrak{U}^{-1} = \mathfrak{E}$

(362) $$\mathfrak{E} = \| \int_{-\infty}^{+\infty} d\tilde{\sigma}_{ik}^{*}(\mu) \| + \sum_{j=1}^{\infty} \| \gamma_i^{(j)} \delta_{ik} \|,$$

wobei die Matrizen rechterhand sowohl in (361) als auch in (362) paarweise getrennt sind. Mit Rücksicht auf das Getrenntsein ist $\| \tilde{R}_{ik}^{*}(\lambda) \|$ offenbar die Resolvente der Matrix

(363) $$\| \tilde{a}_{ik}^{*} \| = \| \int_{-\infty}^{+\infty} \mu \, d\tilde{\sigma}_{ik}^{*}(\mu) \|,$$

und ebenso ist $\left\| \frac{\gamma_i^{(j)} \delta_{ik}}{\lambda - \lambda_j} \right\|$ die Resolvente der Diagonalmatrix $\| \gamma_i^{(j)} \delta_{ik} \lambda_j \|$; entsprechendes gilt auch für die Summe $\sum_{j=1}^{\infty}$ usf. Dies alles ist so zu verstehen, daß man vorher die von der Trennung herrührenden Zeilen und Kolonnen, die mit lauter Nullen besetzt sind, jeweils einfach streicht, bzw. annimmt, daß es überhaupt ein Streckenspektrum gibt. Dem Getrenntsein zufolge setzen sich ferner alle zu $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ gehörigen Eigenlösungen aus Eigenlösungen zusammen, welche von denjenigen Gleichungssystemen herrühren, die zu den *einzelnen* Matrizen

(364) $$\| \tilde{a}_{ik}^{*} \|, \quad \| \gamma_i^{(1)} \lambda_1 \delta_{ik} \|, \quad \| \gamma_i^{(2)} \lambda_2 \delta_{ik} \|, \quad \ldots$$

gehören. Alle diese Matrizen sind, wie etwa aus der Monotonie von $\Phi(\| \sigma_{ik}^{*}(\mu) \|)$ entnommen werden kann, beschränkt und Hermitesch.

§ 92. Nichtexistenz normierter Lösungen in dem Streckenspektrum.

Nehmen wir zunächst an, daß es ein Streckenspektrum gibt. Dann ist $\| \tilde{a}_{ik}^{*} \|$ nicht die Nullmatrix. Denn sonst wäre $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ wegen (361) eine Diagonalmatrix, also wäre die Spektralmatrix von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$, wie aus ihrer Definition mittels der Abschnittsspektralmatrizen hervorgeht, eine

Matrix reiner Sprungfunktionen, es würde also jedes $\tilde{\sigma}_{ik}^{*}(\mu) \equiv 0$ sein, im Widerspruch mit der Voraussetzung. Damit ist zugleich gezeigt, daß $\| a_{ik} \|$ durch kein $\mathfrak{U}$ auf die Diagonalform gebracht werden kann, wenn es ein Streckenspektrum gibt. Gibt es umgekehrt kein Streckenspektrum, so ist jedes $\tilde{\sigma}_{ik}^{*}(\mu) \equiv 0$, da $\hat{\sigma}_{ik}(\mu)$ aus lauter Sprüngen besteht. Also muß dann $\| \tilde{a}_{ik}^{*} \|$ wegen (363) die Nullmatrix sein, so daß $\| a_{ik} \|$ mit Rücksicht auf (361) auf die Diagonalform gebracht werden kann.

Wir beweisen, daß vom Streckenspektrum keine *quadratisch konvergente* Eigenlösung herrühren kann. Ist ein Streckenspektrum vorhanden, so ist $\| \tilde{R}_{ik}^{*}(\lambda) \|$ die Resolvente, also $\| \tilde{\sigma}_{ik}^{*}(\mu) \|$ die Spektralmatrix von $\| \tilde{a}_{ik}^{*} \|$. Hierbei (und fortan immer) ist wieder angenommen, daß die Trennungszeilen und -spalten bereits gestrichen sind, d. h. daß das Schema $\| (i, k) \|$ „kontrahiert" worden ist. Da die $\tilde{\sigma}_{ik}^{*}(\mu)$ durchweg stetig sind, so gilt wegen (313) die Vollständigkeitsrelation

$$(365) \qquad \Big\| \int_{-\infty}^{+\infty} \varphi(\mu)\, d\tilde{\sigma}_{ik}^{*}(\mu) \Big\| \, \Big\| \int_{-\infty}^{+\infty} \psi(\mu)\, d\tilde{\sigma}_{ik}^{*}(\mu) \Big\| = \Big\| \int_{-\infty}^{+\infty} \varphi(\mu)\, \psi(\mu)\, d\tilde{\sigma}_{ik}^{*}(\mu) \Big\|,$$

sobald nur $\varphi(\mu)$ und $\psi(\mu)$ beschränkt und etwa abteilungsweise stetig sind; vgl. S. 192. Freilich brauchen φ und ψ nur auf einem hinreichend großen Intervall beschränkt zu sein. Denn außerhalb eines solchen ist jedes $\tilde{\sigma}_{ik}^{*}(\mu)$ konstant. — Es sei λ eine beliebige reelle Zahl. Es ist

$$(366) \qquad \| \lambda\, \delta_{ik} - \tilde{a}_{ik}^{*} \| = \Big\| \int_{-\infty}^{+\infty} (\lambda - \mu)\, d\tilde{\sigma}_{ik}^{*}(\mu) \Big\|,$$

da wegen (305)

$$(367) \qquad \| \tilde{a}_{ik}^{*} \|^{\nu} = \Big\| \int_{-\infty}^{+\infty} \mu^{\nu}\, d\tilde{\sigma}_{ik}^{*}(\mu) \Big\|; \qquad \nu = 0, 1, \ldots$$

gilt. Setzt man in (365) $\varphi(\mu) \equiv \psi(\mu) \equiv \lambda - \mu$, so folgt für das Quadrat der Matrix (366) die Darstellung

$$(368) \qquad \| A_{ik} \| = \Big\| \int_{-\infty}^{+\infty} (\lambda - \mu)^2\, d\tilde{\sigma}_{ik}^{*}(\mu) \Big\|.$$

Es sei $\mathfrak{y} = (\eta_1, \eta_2, \ldots)$ irgendeine quadratisch konvergente Lösung der zu der Matrix (366) gehörigen homogenen Gleichungen. Dann genügt $\mathfrak{y}$ auch den zu (368) gehörigen homogenen Gleichungen. Denn aus $\mathfrak{B}\,\mathfrak{y} = \mathfrak{o}$ folgt $\mathfrak{B}(\mathfrak{B}\,\mathfrak{y}) = \mathfrak{B}\,\mathfrak{o}$, d. h. $\mathfrak{B}^2\,\mathfrak{y} = \mathfrak{o}$. Man hat also $\sum_{k=1}^{\infty} A_{ik}\,\eta_k = 0$, mithin auch $\sum_{i=1}^{\infty} \bar{\eta}_i \big(\sum_{k=1}^{\infty} A_{ik}\,\eta_k \big) = 0$. Da $\| A_{ik} \|$, als Quadrat einer beschränkten Matrix, beschränkt, und $\{\eta_i\}$ nach Voraussetzung quadratisch konvergent

ist, so kann man hierfür nach (215), S. 125 auch $\sum_{i=1}^{\infty}\sum_{k=1}^{\infty} A_{ik}\eta_i\bar{\eta}_k = 0$ schreiben. Setzt man hierin die Darstellung (368) ein, so folgt

$$\int_{-\infty}^{+\infty} (\lambda-\mu)^2\, d\sum_{i=1}^{\infty}\sum_{k=1}^{\infty}\tilde{\sigma}_{ik}^{*}(\mu)\,\eta_i\bar{\eta}_k = 0.$$

Die Vertauschbarkeit der Grenzübergänge rechtfertigt sich dabei wörtlich so wie schon so oft. Nun ist aber die hinter dem d-Zeichen stehende Funktion, als Kopplungsform einer Spektralmatrix, bei festen η eine nicht abnehmende Funktion von μ. Wegen $(\lambda-\mu)^2 \geqq 0$ muß also auch

$$\int_{-\infty}^{+\infty} d\sum_{i=1}^{\infty}\sum_{k=1}^{\infty}\tilde{\sigma}_{ik}^{*}(\mu)\,\eta_i\bar{\eta}_k = 0$$

gelten, und hier darf man, aus denselben Gründen wie vorher, die Grenzübergänge vertauschen:

$$\sum_{i=1}^{\infty}\sum_{k=1}^{\infty}\eta_i\bar{\eta}_k \int_{-\infty}^{+\infty} d\tilde{\sigma}_{ik}^{*}(\mu) = 0.$$

Die Matrix dieser Form ist aber wegen (367) gleich $\|\tilde{a}_{ik}\|^0 = \|\delta_{ik}\|$, so daß

$$\sum_{i=1}^{\infty}\sum_{k=1}^{\infty}\eta_i\bar{\eta}_k\delta_{ik} = \sum_{i=1}^{\infty}|\eta_i|^2 = 0$$

gilt. Mithin ist jedes $\eta_i = 0$, d. h. $\mathfrak{y}$ ist ein Nullvektor, w. z. b. w.: von der in (361) rechterhand stehenden ersten Matrix (363) rührt — im Sinne der bei (364) gemachten Bemerkung — keine *quadratisch konvergente* Eigenlösung für $\lambda\mathfrak{E} - \mathfrak{A}$ oder $\lambda\mathfrak{E} - \mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ her. Die kursiv gesetzte Einschränkung ist, wie wir sehen werden, sehr wesentlich.

§ 93. Gewicht und Vielfachheit der Eigenwerte.

Wir haben uns noch mit den unter (348) stehenden getrennten Sprungmatrizen $\|\underset{\lambda_j}{\Delta}\sigma_{ik}^{*}\| = \|\gamma_i^{(j)}\delta_{ik}\|$; $j = 1, 2, \ldots$ zu beschäftigen. Alle λ_j sind hierbei, wie erwähnt, voneinander verschieden. Die Anzahl der Einer in der Folge der Zahlen $\gamma_1^{(j)}, \gamma_2^{(j)}, \gamma_3^{(j)}, \ldots$ (die ja entweder $= 0$ oder $= 1$ sind) wollen wir Gewicht von λ_j nennen und vorübergehend mit $N(\lambda_j)$ bezeichnen. $N(\lambda_j)$ ist also die Anzahl der λ_j in der Diagonale der Diagonalmatrix $\|\gamma_i^{(j)}\delta_{ik}\lambda_j\|$, die darüber hinaus höchstens nur Nullen enthalten kann. Es ist gewiß, daß $N(\lambda_j) \neq 0$, d. h. $\geqq 1$ ist. Denn sonst wäre jedes $\underset{\lambda_j}{\Delta}\sigma_{ik}^{*}$; $i, k = 1, 2 \ldots$ gleich Null, d. h. jedes $\sigma_{ik}^{*}(\mu)$ wäre für $\mu = \lambda_j$ stetig, im Widerspruch damit, daß λ_j dem

Punktspektrum angehört. Hingegen kann $N(\lambda_j) = +\infty$ sein (ohne daß dabei notwendig jedes $\gamma_i^{(j)} = 1$ wäre). Der Hilbertsche Satz auf S. 194 wird also bewiesen sein, wenn wir zeigen, daß das Gewicht mit der Vielfachheit identisch (also insbesondere unitär invariant) ist, $N(\lambda_j) = \nu(\lambda_j)$. Zu diesem Zwecke betrachten wir zunächst eine Stelle des Punktspektrums, oder, wie wir jetzt schon sagen wollen, einen Eigenwert, etwa λ_{j_0}. — Es sei λ irgendeine Zahl. Um die zu $\lambda\mathfrak{E} - \mathfrak{A}$ gehörigen quadratisch konvergenten Eigenlösungen aufzufinden, haben wir noch, da in (361) die Matrix $\|\tilde{o}_{ik}^*\|$ bereits erledigt ist, wie erwähnt, die voneinander getrennten unendlichen Gleichungssysteme

$$\sum_{k=1}^{\infty}(\lambda\,\delta_{ik} - \gamma_i^{(j)}\,\delta_{ik}\,\lambda_j)\,x_k^{(j)} = 0\,, \qquad i, j = 1, 2, \ldots,$$

d. h.

$$(369) \qquad \begin{array}{ll} (\lambda - \gamma_i^{(1)}\lambda_1)\,x_i^{(1)} = 0, & i = 1, 2, \ldots; \\ (\lambda - \gamma_i^{(2)}\lambda_2)\,x_i^{(2)} = 0, & i = 1, 2, \ldots; \\ \ldots\ldots\ldots\ldots\ldots & \\ (\lambda - \gamma_i^{(j)}\,\lambda_j)\,x_i^{(j)} = 0, & i = 1, 2, \ldots; \\ \ldots\ldots\ldots\ldots\ldots & \end{array}$$

zu betrachten. — Es sei j_0 irgendwie gewählt. Ist $j \neq j_0$ und $\lambda_{j_0} \neq 0$, so kann der Faktor $(\lambda_{j_0} - \gamma_i^{(j)}\,\lambda_j)$ nicht für alle i verschwinden. Denn mindestens eine der Zahlen $\gamma_1^{(j)}, \gamma_2^{(j)}, \ldots$ ist wegen $N(\lambda_j) \geqq 1$ gleich Eins, und andererseits folgt aus $n \neq m$, wie bereits mehrfach betont wurde, $\lambda_n \neq \lambda_m$ (nach Definition). Mithin kann für $\lambda = \lambda_{j_0}$ das j-te Gleichungssystem in (369) nur durch die Werte $x_i^{(j)} = 0$, $i = 1, 2, \ldots$, befriedigt werden, wenn $j \neq j_0$ ist. Folglich sind für $\lambda = \lambda_{j_0}$ alle Eigenlösungen des gesamten Gleichungssystems (369) unter den Eigenlösungen des einzigen (j_0-ten) Systems

$$(370) \qquad (\lambda - \gamma_i^{(j_0)}\,\lambda_{j_0})\,x_i = 0, \quad i = 1, 2, \ldots; \quad \lambda = \lambda_{j_0},$$

zu suchen. Man kann hierfür, da $\lambda_{j_0} \neq 0$ angenommen ist, auch

$$(371) \qquad (1 - \gamma_i^{(j_0)})\,x_i = 0, \quad i = 1, 2, \ldots$$

schreiben. Die Anzahl der linear unabhängigen quadratisch konvergenten Eigenlösungen dieses Gleichungssystems ist aber gleich der Anzahl derjenigen i, für welche $1 - \gamma_i^{(j_0)} = 0$ ist. Denn ist z. B. $\gamma_1^{(j_0)} = 1$, so kann man (371) durch $x_1 = 1$, $x_2 = 0$, $x_3 = 0$, ... befriedigen; ist aber $\gamma_1^{(j_0)} = 0$, so kann (371) gewiß nur dadurch befriedigt werden, daß man $x_1 = 0$ annimmt, usf. (vgl. S. 27). Nun ist aber die Anzahl derjenigen i, für welche $1 - \gamma_i^{(j_0)} = 0$ ist, nach Definition gleich dem

Gewichte $N(\lambda_{j_0})$ von λ_{j_0}, das sich daher als mit der Vielfachheit $\boldsymbol{v}(\lambda_{j_0})$ identisch erweist. — Es war bisher $\lambda_{j_0} \neq 0$ angenommen, um so aus (370), d. h. aus $\lambda_{j_0}(1-\gamma_i^{(j_0)})\, x_i = 0$, auf (371) schließen zu können. Ist $\lambda_{j_0} = 0$, so ist freilich (370) identisch erfüllt. Es liegt aber — schon mit Rücksicht auf die Trennungsnullen — in der Natur der Sache, daß die Vielfachheit jetzt nicht durch die Lösungen von (370) festgelegt wird, sondern daß auch jetzt $N(\lambda_{j_0}) = \boldsymbol{v}(\lambda_{j_0})$ gesetzt wird. Der auf S. 194 ausgesprochene Hilbertsche Satz ist jedoch auch ohne diese Festsetzung bewiesen, da $\boldsymbol{v}(\lambda_{j_0}) = N(\lambda_{j_0}) \geqq 1$, $\lambda_{j_0} \neq 0$, und, wenn die Festsetzung $\boldsymbol{v}(0) = N(0) \geqq 1$ nicht getroffen wird, $\boldsymbol{v}(0) = +\infty$ gilt.

§ 94. Die Hellingersche Hauptachsentheorie des Streckenspektrums.

Von dem streckenspektralen Bestandteil, $\|\tilde{a}_{ik}^*\|$, wissen wir bisher noch sehr wenig. Wir wissen, daß er [z. B. wegen (367)] die Spektralmatrix $\|\tilde{\sigma}_{ik}^*\|$ besitzt, und daß er nicht auf die Diagonalform gebracht werden kann. Doch besitzt er noch keine Normalform. Denn hat $\|a_{ik}\|$ kein Punktspektrum, so wird $\mathfrak{U}\,\|a_{ik}\|\,\mathfrak{U}^{-1}$ offenbar bei jedem $\mathfrak{U}$, also bereits für $\mathfrak{U} = \mathfrak{E}$, dieselben Eigenschaften wie $\|\tilde{a}_{ik}^*\|$ haben. So müssen wir jetzt nach Hellinger zur Analyse von $\|\tilde{a}_{ik}^*\|$, oder, was dasselbe ist, derjenigen beschränkten Hermiteschen Matrizen übergehen, die kein Punktspektrum haben. — Wir bemerken vorerst, daß in (361) alles, was auf die Hauptdiagonale transformiert werden kann, bereits dorthin gebracht ist. D. h., es ist nicht möglich, eine unitäre Matrix $\mathfrak{W}$ zu finden derart, daß $\mathfrak{W}\,\|\tilde{a}_{ik}^*\|\,\mathfrak{W}^{-1}$ in zwei getrennte Summanden zerfällt, von welchen der eine eine Diagonalmatrix wäre. Es gibt also kein $\mathfrak{W}$ derart, daß die Kopplungsform von $\mathfrak{W}\,\|\tilde{a}_{ik}^*\|\,\mathfrak{W}^{-1}$ in der Gestalt

$$b_{11}\, x_1 \bar{x}_1 + \sum_{i=2}^{\infty} \sum_{k=2}^{\infty} b_{ik}\, x_i \bar{x}_k$$

erscheint, indem $b_{i1} = \bar{b}_{1i} = 0$; $i = 2, 3, \ldots$ gilt (es sei freilich dabei $b_{11} \neq 0$ angenommen). Sonst würde nämlich, wie man z. B. durch Abschnittsbildung und Grenzübergang $(n \longrightarrow +\infty)$ leicht ausrechnet, das (1, 1)-te Element der Spektralmatrix der Matrix $\|b_{ik}\|$ im Punkte $\mu = b_{11}$ den Sprung Eins erleiden, während doch $\|b_{ik}\| = \mathfrak{W}\,\|\tilde{a}_{ik}^*\|\,\mathfrak{W}^{-1}$ kein Punktspektrum haben darf, da $\|\tilde{a}_{ik}^*\|$ kein solches hat. — Die in § 91 betrachtete unitäre Transformation kann also für die Matrizen ohne Punktspektrum wirklich gar nichts leisten.

Nun wissen wir aber (vgl. S. 165 ff.), daß es von μ unabhängige unitäre Transformationen $\mathfrak{W}$ gibt derart, daß

$$(372)\qquad \mathfrak{W}\,\|\tilde{\sigma}_{ik}^{*}(\mu)\|\,\mathfrak{W}^{-1} = \left\| \int_{-\infty}^{\mu} \frac{d\tau_i^{(1)}(\mu)\, d\bar{\tau}_k^{(1)}(\mu)}{d\varrho_1(\mu)} \right\| + \left\| \int_{-\infty}^{\mu} \frac{d\tau_i^{(2)}(\mu)\, d\bar{\tau}_k^{(2)}(\mu)}{d\varrho_2(\mu)} \right\| + \cdots$$

gilt, wobei die Matrizen rechterhand paarweise getrennt sind und die τ und die ϱ passend zu wählende, den Hellingerschen Bedingungen erster Art (S. 108) genügende stetige Funktionen bedeuten; $\varrho_j(\mu)$ ist $\geqq 0$, nicht abnehmend und außerhalb des Streckenspektrums gewiß konstant. Es kann jedoch z. B. $\varrho_1(\mu)$ auch auf einer Teilmenge des Streckenspektrums konstant sein. Die Zerlegung (372) ist nicht eindeutig bestimmt. Man kann nämlich zunächst jede der rechterhand stehenden Matrizen nach den Schlußweisen von § 76 immer weiter in Matrizen von derselben Struktur aufspalten. Doch ist dies nicht die einzige wesentliche Quelle der Möglichkeit von verschiedenen Zerlegungen (372). Es können nämlich „Bestandteile“, die bei einer Zerlegung auf verschiedene Summanden verteilt sind, bei einer auf eine andere Weise, d. h. mit einem anderen $\mathfrak{W}$ angefangenen Zerlegung u. U. nie voneinander getrennt werden, wie weit man dabei auch die Weiterteilung treibt. Doch müssen wir uns hier wieder kurz fassen und uns auf die Wiedergabe der Ergebnisse der Dissertation von Hellinger beschränken.

Wir wissen, daß die Matrix (372) die Spektralmatrix von $\mathfrak{W}\,\|\tilde{a}_{ik}^{*}\|\,\mathfrak{W}^{-1}$ ist. Es folgt daraus wegen (305), wobei $\nu = 1$ zu setzen ist,

$$(373)\qquad \mathfrak{W}\,\|\tilde{a}_{ik}^{*}\|\,\mathfrak{W}^{-1} = \left\| \int_{-\infty}^{+\infty} \mu\, d\xi_{ik}^{(1)}(\mu) \right\| + \left\| \int_{-\infty}^{+\infty} \mu\, d\xi_{ik}^{(2)}(\mu) \right\| + \cdots,$$

mit

$$(374)\qquad \xi_{ik}^{(j)}(\mu) = \int_{-\infty}^{\mu} \frac{d\tau_i^{(j)}\, d\bar{\tau}_k^{(j)}}{d\varrho^{(j)}},$$

und zwar ist wieder $\|\xi_{ik}^{(j)}\|$ die Spektralmatrix der j-ten Matrix in (373) rechterhand, die von der l-ten getrennt ist, wenn $l \neq j$. Die Kopplungsform der j-ten Matrix in (373) rechterhand ist [vgl. (324)]

$$(375)\qquad \int_{-\infty}^{+\infty} \mu \frac{\left| d \sum_{i=1}^{\infty} \tau_i^{(j)}(\mu)\, x_i^{(j)} \right|^2}{d\varrho^{(j)}(\mu)}.$$

Die Kopplungsform von $\mathfrak{W}\,\|\tilde{a}_{ik}^{*}\|\,\mathfrak{W}^{-1}$ ergibt sich durch Summation dieser getrennten Formen in bezug auf j. Die Bestandteile (373) rechterhand sind nun schon Normalbestandteile — eine weitergehende Lösung läßt

ein streckenspektrales „Hauptachsen"-Problem überhaupt nicht zu. So muß (373) als das einzig mögliche Analogon des Satzes betrachtet werden, daß die Kopplungsform einer endlichen Hermiteschen Matrix (oder einer Hermiteschen Matrix ohne Streckenspektrum) mittels einer unitären Transformation auf die Diagonalgestalt $\mu_1 |x^{(1)}|^2 + \mu_2 |x^{(2)}|^2 + \cdots$ gebracht werden kann. Dies gilt, auch wenn die endliche Matrix allgemeiner normal ist, nur brauchen dann die μ_j nicht notwendig reell zu sein. Es scheint, daß auch die Hellingersche Zerlegung (372), (373) sinngemäß auf normale nicht Hermitesche unendliche Matrizen verallgemeinert werden kann, nur müßte man, entsprechend dem Umstande, daß das Spektrum nicht mehr eindimensional zu sein braucht, die Hellingersche Integrationstheorie und die Stieltjessche Darstellungsformel auf zweifache Integrale übertragen, was mehr auf technische als auf grundsätzliche Schwierigkeiten stoßen dürfte. Die Methoden der Hellingerschen Habilitationsschrift, welche die Betrachtung der Abschnitte vermeiden und die unendliche Matrix direkt angreifen, wären dabei nicht nur in Hinblick auf die technischen Schwierigkeiten, sondern auch deshalb vorzuziehen, weil eine unendliche Matrix normal sein kann, ohne daß ihre Abschnitte es sind. Ein unmittelbares, explizites Analogon zu der Stieltjesschen Umkehr*formel* ist freilich nicht möglich, da, im Gegensatz zu den Hermiteschen Spektren, das Spektrum einen zweidimensionalen Bereich voll ausfüllen kann, so daß von den regulären Stellen aus möglicherweise nur der Rand des Spektrums erreichbar ist (es handelt sich jedoch um analytische Funktionen). Eine Klärung der Verhältnisse wäre auch für gewisse Probleme der Potentialtheorie von Wichtigkeit. Vgl. übrigens die auf S. 119 erwähnten Untersuchungen von J. Radon.

§ 95. Exkurs. Bemerkungen über beliebige beschränkte Matrizen.

Die Schursche Kanonisierung dürfte bei sinngemäßer Fragestellung auf beliebige beschränkte Matrizen übertragen werden können, wobei freilich das, was „unterhalb der Diagonale" steht, der Natur der Sache nach ebenso willkürlich bleiben muß, wie in der endlichen „Volterraschen" Matrix

$$\left\| \begin{matrix} 0 & 0 & 0 & & \\ c_{21} & 0 & 0 & \ddots & \\ c_{31} & c_{32} & 0 & \ddots & \ddots \\ \ddots & \ddots & \ddots & 0 & 0 \\ & & \ddots & c_{n\,n-1} & 0 \end{matrix} \right\| + \mathfrak{D},$$

deren sämtliche Eigenwerte verschwinden. Vgl. hierzu die Konstantenabzählung auf S. 21. Das, was unterhalb der Diagonale steht, hängt mit den Eigenlösungen auch bei den endlichen Matrizen nur auf eine lose und für das unitäre Äquivalenzproblem unmittelbar kaum brauchbare Weise zusammen (das Eigenwertsystem ist, wie auf S. 25 besonders hervorgehoben worden ist, nur bei normalen Matrizen das *vollständige* Invariantensystem, enthält also nur bei diesen Matrizen die Lösung des Äquivalenzproblems). — Ein Unterschied gegenüber den endlichen Matrizen besteht darin, daß, wie E. Schmidt und O. Toeplitz gezeigt haben[1]), beschränkte Matrizen $\mathfrak{A}$ existieren derart, daß die zu $\lambda\mathfrak{E} - \mathfrak{A}$ gehörigen homogenen Gleichungen bei keinem λ *irgendwelche* (nicht notwendig normierte) „Eigenlösung" zulassen (im Hermiteschen Falle ist dies nicht möglich; denn gibt es kein Punktspektrum, so gibt es gewiß ein Streckenspektrum und zu diesem gehören Hellingersche Differentiallösungen; vgl. § 97). Doch scheint dieser Unterschied von sekundärer Bedeutung zu sein. Denn in § 63 habe ich bewiesen, daß *jede* beschränkte Matrix mindestens eine „singuläre" Stelle hat, und in diesem singulären Charakter ist wohl das sinngemäße Analogon zu der Existenz von Eigenwerten bei endlichen Matrizen zu erblicken [indem die Richtigkeit oder die Unrichtigkeit des Alternativsatzes (betreffend die Lösbarkeit entweder des homogenen oder des inhomogenen Systems) als eine Frage für sich aufzufassen ist]. Der Existenzbeweis von mindestens einer singulären Stelle ist ja dem geläufigen funktionentheoretischen Beweis des Fundamentalsatzes der Algebra (angewandt auf die Säkulargleichung) unmittelbar nachgebildet. Dieser Standpunkt trägt also weiter als der lösungstheoretische. Für die Lösung des unitären Äquivalenzproblems beliebiger beschränkter Matrizen bedeutet freilich das Schmidt-Toeplitzsche Phänomen ein wesentliches Hindernis. Denn die Übereinstimmung der Spektra ist nur eine notwendige Bedingung für die unitäre Äquivalenz: um also auch hinreichende Bedingungen zu erhalten, muß man die Art der Singularitäten der Matrix *des näheren* beschreiben, und dies ist nach Schmidt und Toeplitz *lösungstheoretisch* offenbar unmöglich.

§ 96. Das streckenspektrale unitäre Äquivalenzproblem.

Wir kehren nunmehr zurück zu den beschränkten Hermiteschen Matrizen ohne Punktspektrum. Die Übereinstimmung der Spektren

---

[1]) Durch Angabe von Beispielen. Die fraglichen Matrizen sind übrigens nicht normal.

zweier solcher Matrizen $\|\tilde{a}_{ik}^*\|$, $\|\tilde{b}_{ik}^*\|$ ist, wie wir wissen, gewiß notwendig für die unitäre Äquivalenz (S. 163). Doch ist diese Bedingung noch nicht hinreichend, ebenso, wie zwei endliche Hermitesche Matrizen noch nicht unitär äquivalent zu sein brauchen, wenn nur ihre Eigenwerte übereinstimmen, ohne daß dies von den beziehentlichen Vielfachheiten bekannt wäre. Daß die Lösung des Problems, ein notwendiges und zugleich hinreichendes Kriterium für die unitäre Äquivalenz von $\|\tilde{a}_{ik}^*\|$, $\|\tilde{b}_{ik}^*\|$ zu finden, ziemlich schwer sein muß, wird bereits dadurch nahegelegt, daß die Hellingersche Zerlegung auf die mannigfachste Weise geschehen kann. Hellinger gelang es, das Problem in seiner Dissertation dennoch zu erledigen. H. Hahn hat das Hellingersche Kriterium im Rahmen seiner auf S. 119 erwähnten Untersuchung neu hergeleitet und zugleich auf eine übersichtlichere Gestalt gebracht. Es ist leider wegen Raummangel ganz unmöglich, auf diese schwierigen Untersuchungen hier des näheren einzugehen.

Wir erwähnen noch, daß die Möglichkeit einer Zerlegung (373) nach Hellinger den folgenden Satz in sich schließt: Jede Hermitesche beschränkte Matrix ohne Punktspektrum kann unitär in die Summe von getrennten Fourierschen beschränkten Matrizen ohne Punktspektrum übergeführt werden. Mithin ist mittels unitärer Substitutionen eine Zerlegung auch in die Summe von getrennten Jacobischen Matrizen möglich. Dieses Ergebnis ist von Toeplitz direkt, d. h. nicht auf dem Umwege über die Spektrumstheorie, sondern durch (eigentliche) Jacobische Transformationen hergeleitet worden. — Offenbar kann daher jede beschränkte Hermitesche Matrix unitär in eine finite Matrix transformiert werden. Denn die Jacobischen Matrizen sind finit, also ist es auch die Summe von getrennten Jacobischen Matrizen, während der eventuelle punktspektrale Bestandteil getrennt in eine Diagonalmatrix übergeführt werden kann, so daß zum Schluß jede Zeile, also auch jede Spalte höchstens drei von Null verschiedene Elemente hat, womit sogar mehr als die Finitheit bewiesen ist. — Es gibt Jacobische Matrizen, die kein Streckenspektrum haben, sie sind jedoch bei Behandlung „vernünftiger“ Kettenbrüche nie vorgekommen; meistens gibt es sogar nur ein Streckenspektrum.

§ 97. Differentiallösungen.

Zum Schluß wollen wir die Hellingerschen „Differentiallösungen“ betrachten. Man setze in (365) $\varphi(\mu) \equiv \mu$, während $\psi(\mu)$ auf die folgende Weise erklärt sein möge: Es ist $\psi(\mu) \equiv \frac{1}{\mu}$, wenn μ auf dem

Intervall $\alpha \leqq \mu \leqq \beta$ liegt, das den Nullpunkt nicht enthalten möge; außerhalb von $[\alpha, \beta]$ sei $\psi(\mu) \equiv 0$. Dann geht (365) in

$$(376) \qquad \| \tilde{a}_{ik}^* \| \left\| \int_\alpha^\beta \frac{d\tilde{\sigma}_{ik}^*(\mu)}{\mu} \right\| = \left\| \int_\alpha^\beta d\tilde{\sigma}_{ik}^*(\mu) \right\|$$

über. Läßt man $[\alpha, \beta]$ auf den Punkt λ zusammenschrumpfen, so bleibt, zunächst nur rein formal,

$$\| \tilde{a}_{ik}^* \| \left\| \frac{d\tilde{\sigma}_{ik}^*(\lambda)}{\lambda} \right\| = \| d\tilde{\sigma}_{ik}^*(\lambda) \|,$$

d. h.

$$(377) \qquad \lambda\, d\tilde{\sigma}_{ik}^* - \sum_{j=1}^{\infty} \tilde{a}_{ij}^*\, d\tilde{\sigma}_{jk}^*(\lambda) = 0,$$

so daß

$$(378) \qquad \sum_{j=1}^{\infty} (\lambda\, \delta_{ij} - \tilde{a}_{ij}^*)\, \xi_j = 0$$

formal durch $\xi_j = d\tilde{\sigma}_{j1}^*(\lambda)$, aber auch durch $\xi_j = d\tilde{\sigma}_{j2}^*(\lambda), \ldots$ befriedigt wird. Gehört nun λ dem Streckenspektrum an, so ist in der Umgebung des Punktes λ nicht jedes $\tilde{\sigma}_{ik}^*(\mu)$ konstant (und umgekehrt). Für die zu $\lambda \mathfrak{E} - \mathfrak{A}$ gehörigen homogenen Gleichungen rühren demnach auch von dem Streckenspektrum „Eigenlösungen" her, nur sind sie eben „Differentiallösungen". Diese zunächst nicht eben überzeugend erscheinende Betrachtung besitzt einen sehr konkreten und für die Anwendungen (sobald es sich nur um „Wellen" irgendwelcher Art handelt) höchst wichtigen Inhalt. Man hat nämlich (zunächst ebenfalls nur formal, da die Funktionen $\tilde{\sigma}_{ik}^*$ vielleicht nicht einmal differentiierbar sind) $\xi_j = d\tilde{\sigma}_{j1}^*(\lambda) = \frac{d\tilde{\sigma}_{j1}^*(\lambda)}{d\lambda} d\lambda, \ldots.$ Kürzt man also in (377) mit $d\lambda$, so bleibt

$$(379) \qquad \sum_{j=1}^{\infty} (\lambda\, \delta_{ij} - \tilde{a}_{ij}^*) \frac{d\tilde{\sigma}_{j1}^*(\lambda)}{d\lambda} = 0, \ldots.$$

Auch jetzt darf man, auch wenn die Funktionen stetig differentiierbar sind, noch *nicht* aufatmen. Denn es ist nicht ausgemacht, daß die Reihen linkerhand konvergent oder sonst irgendwie sinnvoll sind, und auch wenn dies zutreffen sollte, müßte die Gleichung noch immer erst bewiesen werden. Es handelt sich also um eine heuristische Schlußweise, die jedoch sinngemäß genug ist, um meistens zu einem richtigen Endergebnis zu führen. In der Tat kann man leicht die Anwendungen umfassende Bedingungen angeben, unter welchen diese Betrachtung

gewiß legal ist. Es steht freilich von vornherein fest, daß keine der Eigenlösungen $\xi_j = \frac{d\tilde{\sigma}^*_{j1}(\lambda)}{d\lambda}$, $\xi_j = \frac{d\tilde{\sigma}^*_{j2}(\lambda)}{d\lambda}$, ... quadratisch konvergent ist — doch kommen für die meisten Wellenprobleme bereits von vornherein nur diese „langsam abklingenden" Lösungen in Betracht. — Beachtet man noch, daß jede beschränkte Hermitesche Matrix, wie wir vorher gesehen haben, in eine finite übergeführt werden kann, so erkennt man den Zusammenhang mit der auf S. 139 erwähnten Toeplitzschen Untersuchung. Übrigens können Differentiallösungen und quadratisch konvergente Eigenlösungen überlagert auftreten, wenn nämlich λ sowohl dem Punktspektrum als auch dem Streckenspektrum angehört. — Was endlich diejenigen (eventuellen) Punkte λ des Spektrums anlangt, die möglicherweise Häufungsstellen des Punktspektrums sind, ohne zugleich dem Punkt- oder Streckenspektrum anzugehören, so ergibt sich für solche λ nicht einmal eine Differentiallösung. Diese sich an den Hilbertschen Satz (S. 194) anschließende restlose lösungstheoretische Charakterisierung der Spektrumsbestandteile verdankt man ebenfalls Hellinger.

Bei *gewissen* quadratisch konvergenten rechten Seiten können auch die zu $\lambda \mathfrak{E} - \mathfrak{A}$ gehörigen *inhomogenen* Gleichungen, auch wenn λ in dem Streckenspektrum liegt, eine vernünftige, gewissermaßen den Hellingerschen Differentiallösungen entsprechende Lösung haben [trotzdem, wie wir wissen (§ 62), keine *beschränkte* Reziproke zu $\lambda \mathfrak{E} - \mathfrak{A}$ existiert]. Vgl. nämlich die auf S. 100 gemachte Bemerkung über die Cauchyschen Hauptwerte der Resolventenintegrale.

Es versteht sich von selbst, daß $x_i = d\sigma_{i1}(\lambda)$; $x_i = d\sigma_{i2}(\lambda)$, ... auch dann Lösungen der zu $\lambda \mathfrak{E} - \mathfrak{A}$ gehörigen homogenen Gleichungen darstellen, wenn λ ein Eigenwert und sonst nichts ist. Denn die $d\sigma_{ik}(\lambda) = \underset{\lambda}{\Delta} \sigma_{ik}$ sind die Residuen der $R_{ik}(\lambda)$, ebenso wie bei den endlichen Matrizen.

§ 98. Das trigonometrische Momentenproblem.

Eine heuristische Methode, welche die Theorie der beschränkten Hermiteschen Matrizen $\mathfrak{H}$ auf weitere normale Matrizen zu übertragen gestattet, wird durch das Korrespondenzprinzip (S. 181) geliefert: Man betrachte die Matrizen $F(\mathfrak{H})$, wobei $F(\mu)$ irgendeine, für $-\infty < \mu < +\infty$ erklärte, etwa stetige Funktion bezeichnet. Da sie nicht reellwertig zu sein braucht, so kann das durch F entworfene Bild der Klasse der Hermiteschen Matrizen $\mathfrak{H}$ auch nicht Hermitesche Matrizen enthalten. Der Ansatz hat jedoch eine wesentliche Lücke, da er von vornherein

keine Methode zur Beantwortung der Frage liefert, wann eine vorgelegte Matrix bei gegebenem F in der Form $F(\mathfrak{H})$ dargestellt werden kann. Bei gewissen F läßt sich freilich das „Umkehrproblem“ ohne Schwierigkeit erledigen. Ich will dies, unter Zugrundelegung einer Untersuchung von G. Herglotz über das trigonometrische Momentenproblem, im Falle der speziellen konformen Abbildung $F = e^{\mu\sqrt{-1}}$ durchführen. Es ergibt sich so eine der Hilbert-Hellingerschen parallellaufende, völlig geschlossene Spektral- bzw. Transformationstheorie aller unitären Matrizen, so daß der vollständige analytische Apparat einer Darstellungstheorie mit beschränkten Matrizen vorliegt. — Das Abschnittsprinzip dürfte hierbei schwerlich auf eine direkte Weise zum Ziele führen, da die Abschnitte einer unitären Matrix im allgemeinen nicht unitär sind. Jetzt ist also die Heranziehung eines Momentenproblems viel wesentlicher, als bei den Hermiteschen Matrizen (vgl. S. 183). — Bei allgemeinen konformen Abbildungen $F = F(\mu)$ steht hierbei ein Satz von F. Riesz zur Verfügung.

Bei dem trigonometrischen Momentenproblem wird nach einer zwischen 0 und 2π erklärten, reellwertigen, monoton nicht abnehmenden Funktion $\sigma = \sigma(\varphi)$ gefragt, welche die Bedingungsgleichungen

$$\text{(380)} \qquad \frac{1}{2\pi}\int_0^{2\pi} e(n\varphi)\, d\sigma(\varphi) = \alpha_n; \qquad n = 0, \pm 1, \pm 2, \ldots$$

befriedigt, wobei zur Abkürzung $e(x) = e^{x\sqrt{-1}}$ gesetzt ist und $\{\alpha_n\}$ eine gegebene Zahlenfolge bedeutet, die wegen $\overline{e(n\varphi)} = e(-n\varphi)$ die Bedingung $\alpha_n = \bar{\alpha}_{-n}$ erfüllen muß (so daß es genügt, die zu $n \geqq 0$ gehörigen Gleichungen (380) zu befriedigen). Ist σ gegeben, so bezeichnet man die α_n als die (äquidistanten) trigonometrischen Momente von σ. Bei stetig differentiierbarem σ sind, bis auf den Faktor $\sqrt{-1}$, die Momente einfach die auf Laurentsche Weise geschriebenen Fourierkoeffizienten der Ableitung $\sigma'\,(\geqq 0)$.

Das trigonometrische Momentenproblem ist nicht stets lösbar. Das Kriterium der Lösbarkeit ist, wie Herglotz a. a. O. auf eine direkte Weise gezeigt hat, äquivalent mit dem Kriterium des Carathéodory-Toeplitzschen Problems der Theorie der Potenzreihen mit positivem Realteil im Einheitskreis. Dieses Kriterium kann bekanntlich auf die verschiedensten Weisen formuliert werden. Von Carathéodory ist es in geometrischer Gestalt entdeckt worden. In der rechnerischen Sprache der sogenannten Toeplitzschen Matrizen besteht das Kriterium darin, daß die (wegen $\alpha_n = \bar{\alpha}_{-n}$ gewiß Hermitesche) Form

$$\text{(381)} \qquad \sum_{p=0}^{m}\sum_{q=0}^{m} \alpha_{p-q}\,\xi_p\,\bar{\xi}_q$$

nur nichtnegativer Werte fähig ist, wie groß dabei auch m sein mag. M. a. W., die unendliche Matrix

$$(382) \qquad \| \alpha_{(p-1)-(q-1)} \| = \left\| \begin{matrix} \alpha_0 & \alpha_{-1} & \alpha_{-2} & & \\ \alpha_1 & \alpha_0 & \alpha_{-1} & \alpha_{-2} & \\ \alpha_2 & \ddots & \ddots & \ddots & \ddots \\ & \ddots & & & \end{matrix} \right\| \qquad (\alpha_n = \bar{\alpha}_{-n})$$

muß nichtnegativ definit sein (wir haben dabei nicht einfach $\| \alpha_{p-q} \|$ schreiben wollen, da die Zeiger nicht von 1, sondern von 0 an bis $+\infty$ laufen). Dies ist also die notwendige und hinreichende Bedingung für die Lösbarkeit des trigonometrischen Momentenproblems (380).

Da es sich um Stieltjessche Integrationen handelt, können wir annehmen, daß die Lösung $\sigma(\varphi)$ in jedem inneren Punkte des Integrationsbereiches, also für $0 < \varphi < 2\pi$, von links stetig ist. In den Endpunkten ist eine derartige Normierung, wie wir wissen, im allgemeinen unzulässig. Jetzt können wir jedoch annehmen, daß σ auch in dem Punkte $\varphi = 2\pi$ von links stetig ist. Denn die $e(n\varphi)$ sind nach 2π periodisch, so daß es genügt, wenn wir für den Sprung $\sigma(+0) - \sigma(0) (\geqq 0)$ freie Verfügung haben. Die für die Stieltjessche Integration belanglose additive Konstante wollen wir durch die Annahme $\sigma(0) = 0$ festlegen. Unter Lösungen schlechthin verstehen wir nur Lösungen, welche diesen stets erfüllbaren normierenden Bedingungen genügen. Es ist leicht einzusehen, daß es nur eine (derart normierte) Lösung geben kann. M. a. W., es gibt entweder genau eine oder gar keine Lösung, je nachdem die Matrix (382) nichtnegativ definit ist oder nicht. Es seien nämlich σ_1, σ_2 zwei Lösungen von (380), die nicht abnehmend, für $0 < \varphi \leqq 2\pi$ von links stetig und für $\varphi = 0$ gleich Null sind. Setzt man $n = 0$, so folgt, daß die Totalschwankung sowohl bei σ_1 als auch bei σ_2 gleich $2\pi\alpha_0$ ist. Die Totalschwankung ist aber diesmal mit dem zu $\varphi = 2\pi$ gehörigen Funktionswert identisch. Setzt man daher $\eta(\varphi) = \sigma_2(\varphi) - \sigma_1(\varphi)$, so ist $\eta(0) = \eta(2\pi) = 0$. Da nach Voraussetzung

$$\int_0^{2\pi} e(n\varphi)\, d\eta(\varphi) = 0, \qquad n = 0, \pm 1, \ldots$$

gilt, so folgt daraus nach einer partiellen Integration (S. 87), daß

$$\int_0^{2\pi} \eta(\varphi)\, e(n\varphi)\, d\varphi = 0$$

ist für $n \neq 0$. Dies gilt aber wegen $e(0 \cdot \varphi) \equiv 1$ auch für $n = 0$, so daß $\eta(\varphi)$ eine von links stetige, nach 2π periodische Funktion ist, deren alle Fourierkoeffizienten verschwinden. Nach dem Eindeutigkeitssatz der Fourierreihen folgt daraus $\eta \equiv 0$, d. h. $\sigma_1 \equiv \sigma_2$, w. z. b. w.

§ 99. Anwendung auf die unitären Matrizen.

Es seien gegeben eine unitäre Matrix $\mathfrak{U}$ und ein Punkt $\mathfrak{x} = (x_1, x_2, \ldots)$ auf E. Wir behaupten, daß das trigonometrische Momentenproblem (380) lösbar ist, wenn α_n den Wert der Kopplungsform von $\mathfrak{U}^n$ im Punkte $\mathfrak{x}$ bezeichnet,

$$\alpha_n = \Phi(\mathfrak{U}^n; \mathfrak{x}); \qquad n = 0, \pm 1, \pm 2, \ldots \tag{383}$$

(die nicht abnehmende Lösung $\sigma = \sigma(\varphi)$ hängt dabei außer von $\mathfrak{U}$ auch noch von $\mathfrak{x}$ ab, so daß

$$\sigma = \sigma(\varphi; x_1, x_2, \ldots) \tag{384}$$

ist). Zunächst gilt nämlich wegen $\mathfrak{U}^* = \mathfrak{U}^{-1}$

$$\begin{aligned} \overline{\Phi(\mathfrak{U}; \mathfrak{x})} &= \overline{\sum_{i=1}^{\infty} \sum_{k=1}^{\infty} u_{ik} x_i \bar{x}_k} = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \bar{u}_{ik} \bar{x}_i x_k \\ &= \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \bar{u}_{ki} x_i \bar{x}_k = \Phi(\mathfrak{U}^*; \mathfrak{x}) = \Phi(\mathfrak{U}^{-1}; \mathfrak{x}), \end{aligned} \tag{385}$$

also, da zugleich mit $\mathfrak{U}$ auch $\mathfrak{U}^{-n}$ unitär ist, allgemeiner $\overline{\Phi(\mathfrak{U}^n; \mathfrak{x})} = \Phi(\mathfrak{U}^{-n}; \mathfrak{x})$, so daß durch die Folge (383) die Bedingung $\alpha_n = \bar{\alpha}_{-n}$ gewiß erfüllt wird. Ferner ist jetzt, wenn $\mathfrak{V} = \sum_{p=0}^{m} \xi_p \mathfrak{U}^p$ gesetzt wird, die Form (381) wegen $\mathfrak{U}^{-q} = \mathfrak{U}^{*q}$ gleich

$$\sum_{p=0}^{m} \sum_{q=0}^{m} \alpha_{p-q} \xi_p \bar{\xi}_q = \sum_{p=0}^{m} \sum_{q=0}^{m} \Phi(\mathfrak{U}^{p-q}; \mathfrak{x}) \xi_p \bar{\xi}_q$$

$$= \sum_{p=0}^{m} \sum_{q=0}^{m} \Phi(\xi_p \bar{\xi}_q \mathfrak{U}^p \mathfrak{U}^{-q}; \mathfrak{x}) = \Phi\Big(\Big[\sum_{p=0}^{m} \xi_p \mathfrak{U}^p\Big]\Big[\sum_{q=0}^{m} \bar{\xi}_q \mathfrak{U}^{-q}\Big]; \mathfrak{x}\Big) = \Phi(\mathfrak{V}\mathfrak{V}^*; \mathfrak{x}),$$

sie hat also stets einen Wert, der von der Kopplungsform einer Norm ($= \mathfrak{V}\mathfrak{V}^*$) angenommen wird, ist also nach S. 132 gewiß $\geqq 0$, so daß die Herglotzsche Bedingung erfüllt ist, w. z. b. w.

Setzt man in der Lösung (384) außer x_i und x_k die übrigen x gleich Null, so entsteht eine Funktion, die wir mit $\sigma(x_i, x_k)$ bezeichnen wollen. Sie hängt auch von φ ab. Da es bei jedem $\mathfrak{x}$ auf E oder auch bei jedem quadratisch konvergenten $\mathfrak{x}$ eine nicht abnehmende

Lösung σ gibt, und da diese Lösung, wie wir auf S. 210 gesehen haben, eindeutig bestimmt ist, so muß $\sigma(x_i, x_k)$, wie der Leser selbst näher begründen möge, eine Hermitesche Form sein,

$$\sigma(x_i, x_k) = \sigma_{ii} x_i \bar{x}_i + \sigma_{ik} x_i \bar{x}_k + \sigma_{ki} \bar{x}_i x_k + \sigma_{kk} x_k \bar{x}_k; \quad \sigma_{ik} = \bar{\sigma}_{ki}, \quad i \gtrless k.$$

Ebenso ist die Lösung (384) für den Vektor $(x_1, x_2, \ldots x_N, 0, 0, \ldots)$, den wir mit $\mathfrak{x}_{[N]}$ bezeichnen wollen, gleich

$$(386) \qquad \sigma(\varphi; \mathfrak{x}_{[N]}) = \sum_{i=1}^{N} \sum_{k=1}^{N} \sigma_{ik} x_i \bar{x}_k,$$

wobei die σ_{ik} mit Rücksicht auf die Eindeutigkeit der Lösung die vorige Bedeutung haben, also der Hermiteschen Bedingung $\sigma_{ik} = \bar{\sigma}_{ki}$ genügende Funktionen von φ sind. Es ist dabei $\alpha_n = \Phi(\mathfrak{U}^n; \mathfrak{x}_{[N]})$ zu setzen, so daß

$$(387) \quad \frac{1}{2\pi}\int_0^{2\pi} e(n\varphi)\, d\sigma(\varphi; \mathfrak{x}_{[N]}) = \Phi(\mathfrak{U}^n; \mathfrak{x}_{[N]}); \qquad n = 0, \pm 1, \ldots,$$

also wegen $d\sigma \geqq 0$ insbesondere

$$(388) \qquad 0 \leqq \frac{1}{2\pi}\int_0^{2\pi} d\sigma(\varphi; \mathfrak{x}_{[N]}) = \Phi(\mathfrak{U}^0; \mathfrak{x}_{[N]}) = \sum_{i=1}^{N} |x_i|^2 \leqq 1$$

ist. Da σ für $\varphi = 0$ verschwindet und mit wachsendem φ nicht abnimmt, so folgt hieraus wegen (386), daß auf E .

$$(389) \qquad 0 \leqq \sum_{i=1}^{N} \sum_{k=1}^{N} \sigma_{ik}(\varphi)\, x_i \bar{x}_k \leqq 2\pi$$

gilt, so daß die Matrix $\|\sigma_{ik}(\varphi)\|$, die wir, nachdem sie nötigenfalls linksseitig stetig gemacht worden ist, als die *Spektralmatrix der unitären Matrix* $\mathfrak{U}$ bezeichnen wollen, beschränkt ist. In jedem festen Punkte von E ist $\Phi(\|\sigma_{ik}(\varphi)\|; \mathfrak{x})$ eine nicht abnehmende, für $0 \leqq \varphi \leqq 2\pi$ erklärte Funktion, die für $\varphi = 0$ verschwindet und für $\varphi = 2\pi$ gleich 2π ist [vgl. (380)]. — Nehmen wir in (387) bei festem n den Grenzübergang $N \to +\infty$ vor, so konvergiert die rechte Seite nach (217) gegen $\Phi(\mathfrak{U}^n; \mathfrak{x})$, da $\mathfrak{U}^n$ unitär, also beschränkt ist. Was die linke Seite von (387) betrifft, so ist es gestattet, den Grenzübergang $N \to +\infty$ einfach unter dem Integralzeichen vorzunehmen. Denn die Bedingungen des Hellyschen Satzes über gliedweise Integration sind nach dem soeben Bewiesenen gewiß erfüllt. Nun gilt aber wegen (386) die Grenzgleichung

$\sigma(\varphi; \mathfrak{x}_{[N]}) \to \Phi(\|\sigma_{ik}(\varphi)\|; \mathfrak{x})$, so daß also

$$(390)\quad \frac{1}{2\pi}\int_0^{2\pi} e(n\varphi)\, d\Phi(\|\sigma_{ik}(\varphi)\|; \mathfrak{x}) = \Phi(\mathfrak{U}^n; \mathfrak{x}); \qquad n = 0, \pm 1, \ldots$$

ausfällt. Die Herleitung dieser Formel war unser Ziel; sie besagt, daß die Kopplungsform der Spektralmatrix von $\mathfrak{U}$ identisch ist mit der Lösung (384) unseres Momentenproblems (380), (383).

In der Matrizensprache lautet (390) wie folgt:

$$(391)\quad \left\| \frac{1}{2\pi}\int_0^{2\pi} e(n\varphi)\, d\sigma_{ik}(\varphi) \right\| = \mathfrak{U}^n; \qquad n = 0, \pm 1, \ldots,$$

woraus ersichtlich ist (vgl. S. 183), warum wir das Momentenproblem (380), (383) gestellt haben (über die Frage, wie man eben auf dieses Momentenproblem geführt wird, gibt die Schlußbemerkung auf S. 217 Aufschluß).

§ 100. Die Resolvente der unitären Matrizen.

Da $\mathfrak{U}$ unitär ist, so ist wegen (252) offenbar $\mathsf{P}(\mathfrak{U}) = 1$, und daher, da auch $\mathfrak{U}^n$ unitär ist, $\mathsf{P}(\mathfrak{U}^n) = 1$, $n = 0, \pm 1, \ldots$, so daß wegen (230) die Matrizenreihe

$$(392)\quad \sum_{n=0}^{\infty} \frac{\mathfrak{U}^n}{\lambda^{n+1}} \quad \text{für} \quad |\lambda| > 1,$$

und ebenso

$$(393)\quad -\sum_{n=0}^{\infty} \lambda^n \mathfrak{U}^{-n-1} \quad \text{für} \quad |\lambda| < 1$$

gewiß konvergiert. Sie stellen (vgl. S. 144) eine beschränkte Reziproke der beschränkten und normalen Matrix $\lambda\mathfrak{E} - \mathfrak{U}$ für $|\lambda| > 1$ bzw. $|\lambda| < 1$ dar, so daß alle singulären Stellen der unitären Matrizen auf dem Rande $|\lambda| = 1$ des Einheitskreises liegen müssen. Mindestens eine singuläre Stelle gibt es dabei nach S. 145 gewiß (für $\mathfrak{U} = \mathfrak{E}$ gibt es nur eine einzige). Aus (391) folgt offenbar (vgl. S. 179)

$$\sum_0^{\infty} \frac{\mathfrak{U}^n}{\lambda^{n+1}} = \left\| \frac{1}{2\pi}\int_0^{2\pi} \sum_{n=0}^{\infty} e(n\varphi) \frac{d\sigma_{ik}(\varphi)}{\lambda^{n+1}} \right\|$$

$$= \left\| \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{\lambda} \sum_{n=0}^{\infty} \left[\frac{e(\varphi)}{\lambda}\right]^n d\sigma_{ik}(\varphi) \right\| = \left\| \frac{1}{2\pi}\int_0^{2\pi} \frac{d\sigma_{ik}(\varphi)}{\lambda - e(\varphi)} \right\|,$$

so daß, wenn wieder $(\lambda \mathfrak{E} - \mathfrak{U})^{-1} = \| R_{ik}(\lambda) \|$ gesetzt wird,

$$R_{ik}(\lambda) = \frac{1}{2\pi} \int_0^{2\pi} \frac{d\sigma_{ik}(\varphi)}{\lambda - e(\varphi)} \tag{394}$$

ausfällt. Dies gilt zunächst nur unter der Voraussetzung $|\lambda| > 1$. Da die beiden Gebiete $|\lambda| > 1$, $|\lambda| < 1$ durch die geschlossene Kurve $|\lambda| = 1$ getrennt werden, so wäre es denkbar, daß (394) für $|\lambda| < 1$ nicht mehr gilt. Daß dem nicht so ist, folgt durch Einsetzen von (391) diesmal in (393). Folglich gilt (394) für $\lambda \neq 1$ gewiß.

Unter Spektrum von $\mathfrak{U}$ wollen wir die Menge derjenigen Punkte der Kreislinie $|\lambda| = 1$ oder $\lambda = e(\varphi)$ verstehen, die Nichtkonstanzstellen von mindestens einem $\sigma_{ik}(\varphi)$ oder von Häufungsstellen von solchen Nichtkonstanzstellen sind, so daß φ_0 dann und nur dann außerhalb des Spektrums liegt, wenn es ein $\varepsilon > 0$ gibt derart, daß für $\varphi_0 - \varepsilon < \varphi < \varphi_0 + \varepsilon$ alle $\sigma_{ik}(\varphi)$ konstant sind. Ist $\{\varphi_i\}$ eine beliebige reelle Zahlenfolge, also $\{e(\varphi_i)\}$ eine Punktfolge auf $|\lambda| = 1$, so ist die Diagonalmatrix $\| e(\varphi_i)\, \delta_{ik} \|$ offenbar unitär[1]), und ihre singulären Stellen werden durch diese Punktfolge und ihre Häufungsstellen gebildet, so daß jede auf $|\lambda| = 1$ gelegene abgeschlossene Punktmenge das Spektrum einer unitären Matrix bilden kann. Es kann also z. B. aus einem oder mehreren getrennten Bögen bestehen und auch die volle Kreislinie bedecken. Gibt es aber einen Punkt $e(\varphi_0)$ auf $|\lambda| = 1$, der nicht dem Spektrum angehört, so ist $\lambda = e(\varphi_0)$ eine matrizentheoretisch reguläre Stelle von $\mathfrak{U}$, und (394) gilt auch für $\lambda = e(\varphi_0)$, trotzdem $|\lambda| = 1$ ist. Man erkennt dies durch wortgetreue Wiederholung der auf S. 178 angewandten Schlußweise. Außerhalb des Spektrums ist also $\mathfrak{U}$ gewiß regulär und es gilt dabei (394). Die durch die Formel (394) dargestellten Funktionen sind in jedem außerhalb des Spektrums gelegenen Bereiche regulär. Wir werden sogleich beweisen, daß in einem Gebiete, das Punkte des Spektrums enthält, mindestens eine der durch (394) dargestellten Funktionen $R_{ik}(\lambda)$ singulär ist. Erst recht ist also jeder Punkt des Spektrums eine singuläre Stelle von $\mathfrak{U}$ (singulär in dem auf S. 142 erklärten matrizentheoretischen Sinne; vgl. S. 152). Einerseits ist also das Spektrum identisch mit der Gesamtheit der matrizentheoretisch singulären Stellen. Andererseits ist eine Stelle dann und nur dann matrizentheoretisch singulär, wenn sie eine singuläre Stelle mindestens einer durch (394) dargestellten Funktion oder ein Häufungs-

[1]) Und hat nur ein Punktspektrum. Die Definitionen auf S. 177 sollen auch für unitäre Matrizen in Anspruch genommen werden.

punkt von solchen Stellen ist. Wir sprechen dabei, und zwar aus ähnlichen Gründen wie auf S. 99, nicht von der analytischen Funktion $R_{ik}(\lambda)$, sondern von der durch (394) unmittelbar dargestellten Funktion.

Der Beweis für die erwähnte Kennzeichnung der singulären Stellen eines Integrals

$$R(\lambda) = \frac{1}{2\pi}\int_0^{2\pi} \frac{d\tau(\varphi)}{\lambda - e(\varphi)}, \tag{395}$$

wobei $\tau(\varphi)$ von beschränkter Schwankung ist, gestaltet sich wie folgt. Es liegt mit Rücksicht auf S. 96 nahe, die Differenz $R(\lambda) - R(\bar\lambda^{-1})$ zu bilden, da $\bar\lambda^{-1}$ im Sinne der Abbildung durch reziproke Radien das Spiegelbild von λ in bezug auf den Einheitskreis ist. Man findet, wenn etwa $|\lambda| < 1$ angenommen und $\lambda = |\lambda|\, e(\theta)$ gesetzt wird,

$$\frac{1}{\lambda - e(\varphi)} - \frac{1}{\bar\lambda^{-1} - e(\varphi)} = \frac{\bar\lambda^{-1} - \lambda}{\lambda\bar\lambda^{-1} - (\lambda + \bar\lambda^{-1})\, e(\varphi) + e(2\varphi)}$$

$$= e(-\varphi)\frac{1 - \lambda\bar\lambda}{\lambda\, e(-\varphi) - (\lambda\bar\lambda + 1) + \bar\lambda\, e(\varphi)} = -\, e(-\varphi)\frac{1 - |\lambda|^2}{|\lambda|^2 + 1 - 2|\lambda|\cos(\varphi - \theta)}.$$

Setzt man also

$$\varrho(\varphi) = -\int e(\varphi)\, d\tau(\varphi), \tag{396}$$

so folgt nach § 40

$$R(\lambda) - R\left(\frac{1}{\bar\lambda}\right) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1 - |\lambda|^2}{1 - 2|\lambda|\cos(\varphi - \theta) + |\lambda|^2}\, d\varrho(\varphi) \quad (|\lambda| \neq 1). \tag{397}$$

Es sei nun $R(\lambda)$ in jedem Punkte eines Bogens $\theta_0 - \varepsilon \leqq \theta \leqq \theta_0 + \varepsilon\ (>\theta_0)$ der Kreislinie $\lambda = e(\theta)$, wobei also $|\lambda| = 1$ ist, regulär. Dann ist $R(\lambda)$ gewiß gleichmäßig stetig in einem Gebiete, das den Bogen ganz in seinem Innern enthält, also muß $R(\lambda)$ auf den beiden Seiten ($|\lambda| < 1$ bzw. $|\lambda| > 1$) des Bogens stetig in die Werte übergehen, die $R(\lambda)$ auf dem Bogen hat. Konvergiert ferner der Punkt $|\lambda|\, e^{i\theta}$ gegen $e^{i\theta}$, $|\lambda| \to 1 - 0$, so hat dabei sein Spiegelbild $\bar\lambda^{-1} = |\lambda|^{-1} e^{i\theta}$ dieselbe Grenzlage $e^{i\theta}$, so daß auf dem Bogen gleichmäßig die Limesgleichung $R(\lambda) - R(\bar\lambda^{-1}) \to 0$, $|\lambda| \to 1 - 0$ gilt. Mit Rücksicht auf (397) folgt daraus nach neueren Untersuchungen über diese von F. Riesz eingeführten Stieltjes-Poissonschen Integrale, daß die Belegung $\varrho(\varphi)$ für $\theta_0 - \varepsilon < \theta < \theta_0 + \varepsilon$ konstant ist. Eine partielle Integration in (396) ergibt alsdann, daß dabei auch $\tau(\varphi)$ konstant sein muß, w. z. b. w. Sind

freilich die $\sigma_{ik}(\varphi)$ der Spektralmatrix stetig differentiierbar, so ist die vorhin gegebene spektrale Charakterisierung der singulären Stellen von $\mathfrak{U}$ eine unmittelbare Folgerung aus dem klassischen Ergebnis von H. A. Schwarz, das aus den Elementen der Theorie des logarithmischen Potentials bzw. der Fourierreihen bekannt ist.

§ 101. Spektraltheorie der unitären Matrizen.

Ist $f(\varphi)$ eine für $0 \leqq \varphi \leqq 2\pi$ erklärte und etwa stetige Funktion, so wollen wir

$$(398) \qquad f(\mathfrak{U}) = \left\| \frac{1}{2\pi} \int_0^{2\pi} f(\varphi)\, d\sigma_{ik}(\varphi) \right\|$$

setzen. Man überzeugt sich leicht (vgl. S. 182), daß diese Matrix beschränkt ist. Es sei $g(\varphi)$ eine andere, ebenso wie f beschaffene Funktion. Zunächst mögen f und g abbrechende, nach 2π periodische Fourierreihen (trigonometrische Polynome) sein. Aus (391) folgt dann, ebenso wie auf S. 181, die Vollständigkeitsrelation

$$(399) \qquad f(\mathfrak{U})\, g(\mathfrak{U}) = (f g)(\mathfrak{U}),$$

unter der Matrix rechterhand die zu der Funktion $h(\varphi) = f(\varphi)\, g(\varphi)$ gehörige Matrix $h(\mathfrak{U})$ verstanden. Nachträglich kann man die Bedingung, daß f und g trigonometrische Polynome sind, nach einer früher angedeuteten Schlußweise fallen lassen. — Es gilt ferner

$$f(\mathfrak{U}) + g(\mathfrak{U}) = (f + g)(\mathfrak{U}),$$

so daß wir damit ein Korrespondenzprinzip für unitäre Argumente haben (auf S. 181 war die zugrunde gelegte Matrix Hermitesch). — Entsprechend könnten wir „Fouriersche" unitäre Matrizen betrachten, indem wir auf S. 189 die Funktion $\chi_k(\varphi) = e(k\varphi)$ setzen.

Wir setzen in (398) $f(\varphi) = \varphi^m$; $m = 0, 1, 2, \ldots$. Damit die so erklärte beschränkte und wegen $\sigma_{ik} = \bar{\sigma}_{ki}$ gewiß Hermitesche Matrix nicht mit $\varphi^m(\mathfrak{U})$ bezeichnet werden muß, wollen wir

$$(400) \qquad \left\| \frac{1}{2\pi} \int_0^{2\pi} \varphi^m\, d\sigma_{ik}(\varphi) \right\| = \mathfrak{H}_m; \qquad m = 0, 1, \ldots$$

und insbesondere

$$(401) \qquad \left\| \frac{1}{2\pi} \int_0^{2\pi} \varphi\, d\sigma_{ik}(\varphi) \right\| = \mathfrak{H}_1 = \mathfrak{H}$$

setzen. Jeder unitären Matrix $\mathfrak{U}$ ist damit mittels ihrer Spektralmatrix eindeutig eine beschränkte Hermitesche Matrix $\mathfrak{H}$ zugeordnet. Aus (399) folgt leicht, daß $\mathfrak{H}_m$ die m-te Potenz $\mathfrak{H}^m$ von $\mathfrak{H}$ ist; $m = 0, 1, 2, \ldots$. Die C. Neumannsche Reihe von $\mathfrak{H}$ ist daher wegen (400)

$$(402) \qquad (\lambda\,\mathfrak{E} - \mathfrak{H})^{-1} = \sum_0^\infty \frac{\mathfrak{H}_m}{\lambda^{m+1}} = \left\| \frac{1}{2\pi} \int_0^{2\pi} \frac{d\,\sigma_{ik}(\varphi)}{\lambda - \varphi} \right\|.$$

Führen wir also eine Funktion $\sigma_{ik}(\mu)$ ein, so daß $\sigma_{ik}(\mu) = \sigma_{ik}(\varphi)$ ist, wenn $0 \leqq \varphi = \mu \leqq 2\pi$ ausfällt, und daß $\sigma_{ik}(\mu)$, im Gegensatz zu $\sigma_{ik}(\varphi)$, auch für $\mu < 0$ und $\mu > 2\pi$ erklärt wird, und zwar so, daß dabei $\sigma_{ik}(\mu) = 0$ bzw. $\sigma_{ik}(\mu) = \delta_{ik}$ ist, so stellt

$$(403) \qquad \left\| \frac{\sigma_{ik}(\mu)}{2\pi} \right\|$$

mit Rücksicht auf die Stieltjessche Umkehrformel die Spektralmatrix von $\mathfrak{H}$ dar, wenn $\| \sigma_{ik}(\varphi) \|$ die Spektralmatrix von $\mathfrak{U}$ bezeichnet. Da die unitäre Spektralmatrix für $0 < \varphi \leqq 2\pi$ von links stetig ist, so ist $\mu = 2\pi$ offenbar eine Stetigkeitsstelle aller Elemente der (auch für $\mu > 2\pi$ erklärten) Hermiteschen Spektralmatrix, d. h. $\mu = 2\pi$ liegt nicht in dem Punktspektrum von $\mathfrak{H}$. — Ebenso beweist man, daß, wenn $\mathfrak{H}$ eine vorgegebene Hermitesche Matrix bezeichnet, *deren ganzes Spektrum irgendwo auf dem Intervall* $0 \leqq \mu \leqq 2\pi$ *und dabei* $\mu = 2\pi$ *außerhalb des Punktspektrums liegt,* so daß die Spektralmatrix $\| \sigma_{ik}(\mu) \|$ von $\mathfrak{H}$ für $\mu < 0$ und $\mu > 2\pi$ konstant $= \mathfrak{O}$ bzw. $= \mathfrak{E}$ ist, so kann folgendes behauptet werden: Bildet man die Matrix (vgl. S. 183)

$$(404) \qquad \frac{1}{2\pi} e(\mathfrak{H}) = \left\| \frac{1}{2\pi} \int_0^{2\pi} e(\varphi)\, d\sigma_{ik}(\varphi) \right\|,$$

und bezeichnet sie mit $\mathfrak{U}$, so ist $\mathfrak{U}$ aus naheliegenden Gründen unitär (dies gilt, auch wenn das Spektrum von $\mathfrak{H}$ beliebig *breit* ist[1]); vgl. S. 270); und ihre Spektralmatrix ist $\| \sigma_{ik}(\varphi) \|$; $0 \leqq \varphi \leqq 2\pi$. Um dies einzusehen, braucht man jetzt nur umgekehrt von der Hermiteschen Vollständigkeitsrelation (313) auszugehen. Da die auf S. 215 mit $\mathfrak{H}_1$ bezeichnete Matrix mit der unitären Matrix $\mathfrak{U}$, aus welcher $\mathfrak{H}_1$ ab-

[1]) Es kommt für die Eineindeutigkeit der Abbildung nur die Breite, nicht aber auch die Lage des Spektrums von $\mathfrak{H}$ auf der μ-Geraden in Betracht. Denn einerseits sind die Spektra von $\mathfrak{H}$ und $\mathfrak{H} + C\mathfrak{E}$ kongruente Punktmengen, andererseits hat $e(\mathfrak{A})$ die Periode $2\pi\,\mathfrak{E}$.

geleitet wurde, wegen (398) in der Beziehung $\mathfrak{U} = \frac{1}{2\pi} e(\mathfrak{H}_1)$ steht, so sehen wir insbesondere, daß durch die Abbildung $\mathfrak{U} = \frac{1}{2\pi} e(\mathfrak{H})$ die Gesamtheit derjenigen Hermiteschen Matrizen, *deren Spektrum zwischen* 0 *und* 2π *liegt und deren Punktspektrum den Punkt* $\mu = 2\pi$ *nicht enthält,* auf die Gesamtheit aller unitären Matrizen eineindeutig bezogen wird. Dies besagt wesentlich mehr, als daß jede unitäre Matrix in der Form $e(\mathfrak{H})$ mit einem passend gewählten $\mathfrak{H}$ dargestellt werden kann.

Die Hilbert-Hellingerschen Untersuchungen über die verschiedenen Arten von Spektralbestandteilen, die Transformationstheorie auf die Hauptachsen bzw. — wenn es ein Streckenspektrum gibt — die Trennungstheorie usf. können ohne Schwierigkeit auf unitäre Matrizen übertragen werden. Doch soll dies dem Leser überlassen bleiben.

Es ist unnötig zu bemerken, daß die vorhergehenden Betrachtungen auch im Bereich der endlichen Matrizen gelten. Freilich handelt es sich dabei nur um leicht verifizierbare Identitäten, ebenso wie auf S. 49 ff.

Der formal-algebraische Kern dieser ganzen Spektraltheorie der unitären Matrizen ist offenbar in der folgenden Bemerkung enthalten. Ist $\mathfrak{H}_{[n]}$ eine Hermitesche Diagonalmatrix $\|\theta_i \delta_{ik}\|_{[n]}$, so ist $\mathfrak{U}_{(n)} = \|e(\theta_i)\delta_{ik}\|_{(n)}$ eine unitäre Diagonalmatrix und umgekehrt, und zwar gilt $e(\mathfrak{H}_{[n]}) = \mathfrak{U}_{(n)}$. Nun ist aber (vgl. § 7)

$$\mathfrak{W}_{(n)}\, e(\mathfrak{H}_{[n]})\, \mathfrak{W}^*_{(n)} = \mathfrak{W}_{(n)} \sum_{m=0}^{\infty} \frac{\left(\mathfrak{H}_{[n]}\sqrt{-1}\right)^m}{m!}\, \mathfrak{W}^*_{(n)} = e\left(\mathfrak{W}_{(n)}\, \mathfrak{H}_{[n]}\, \mathfrak{W}^*_{(n)}\right),$$

und bei den endlichen unitären oder Hermiteschen Matrizen läßt sich die Diagonalform mittels einer unitären Transformation $\mathfrak{W}_{(n)}$ stets erreichen. [Es ist dabei zu beachten, daß $e(\mathfrak{H}_{[n]})$ nicht der n-te Abschnitt von $e(\mathfrak{H})$ ist, der, wie bereits betont wurde überhaupt keine unitäre Matrix zu sein braucht.]

Sechstes Kapitel.

Hermitesche nicht beschränkte Matrizen.

§ 102. Matrizen mit einer Lücke in dem Abschnittsspektrum.

Es sei $\mathfrak{A}$ eine Matrix, die weder Hermitesch noch beschränkt, ja nicht einmal eine Q-Matrix zu sein braucht. Hingegen sei jeder Abschnitt $\mathfrak{A}_{[n]}$ normal, und es möge die folgende Lückenbedingung erfüllt sein: Es gibt in der komplexen Ebene irgendwo einen Punkt λ_* derart, daß die Abschnittseigenwerte $\lambda_\nu^{(n)}$; $\nu = 1, 2, \ldots, n$, $n = 1, 2, \ldots$, wobei also $\det(\mathfrak{E}_{[n]} \lambda_\nu^{(n)} - \mathfrak{A}_{[n]}) = 0$ gilt, dem Punkte λ_* nicht beliebig nahe kommen; es gibt demnach ein $\varepsilon > 0$ derart, daß das Kreisgebiet $|\lambda - \lambda_*| < \varepsilon$, das wir auch mit Ω bezeichnen wollen, von den $\lambda_\nu^{(n)}$ frei ist. Alle n^2 Elemente $R_{ik}^{(n)}(\lambda)$ der Resolvente von $\mathfrak{A}_{[n]}$ sind daher auf Ω regulär. — Wir behaupten folgendes.

1. Die Folge $\{n\}$ der natürlichen Zahlen enthält eine von i, k und λ unabhängige Teilfolge $\{h_n\}$ derart, daß $\lim R_{ik}^{(h_n)}(\lambda)$ für $n \to \infty$ auf Ω bei jedem i und k existiert und eine reguläre Funktion darstellt. — 2. Wie die Folge $\{h_n\}$ auch sonst beschaffen sei, stellt $\| R_{ik}(\lambda) \| = \| \lim R_{ik}^{(h_n)}(\lambda) \|$ eine beschränkte Matrix dar, die übrigens auf jedem ganz im Innern von Ω gelegenen Gebiet Ω' gleichmäßig beschränkt ist, und zwar gilt gleichmäßig auf Ω'

$$\sum_{k=1}^{h_n} R_{ik}^{(h_n)}(\lambda)\, c_k \to \sum_{k=1}^{\infty} R_{ik}(\lambda)\, c_k, \qquad n \to +\infty \tag{405}$$

für alle i bei einer jeden quadratisch konvergenten Folge $\{c_k\}$. — 3. Ist außerdem $\mathfrak{A}$ eine Q-Matrix, so ist $\| R_{ik}(\lambda) \|$ eine sowohl hintere als auch vordere Reziproke von $\lambda \mathfrak{E} - \mathfrak{A}$, und $x_i = \sum_{k=1}^{\infty} R_{ik}(\lambda)\, c_k$ ist eine

quadratisch konvergente Lösung von $(\lambda\mathfrak{E}-\mathfrak{A})\mathfrak{x}=\mathfrak{c}$ (sobald λ im Lückenkreise liegt).

Dies alles wird bewiesen sein, wenn wir zeigen können, daß es eine nur von $\mathfrak{A}$ und von Ω' abhängige, also von n und von λ unabhängige Schranke $K=K(\Omega',\mathfrak{A})$ gibt derart, daß auf Ω'

(406) $$\mathsf{M}(\| R_{ik}^{(n)}(\lambda)\|_{(n)}) < K(\Omega',\mathfrak{A})$$

gilt (wobei der Ausdruck linkerhand selbstverständlich das Maximum von $|\sum_{i=1}^{n}\sum_{k=1}^{n} R_{ik}^{(n)}(\lambda)x_i\bar{x}_k|$ auf $\mathsf{E}_{[n]}$ bezeichnet). Dann können nämlich die Schlüsse von S. 169 ff. wörtlich wiederholt werden.

Ist $\mathfrak{A}$ Hermitesch, so ist die Lückenbedingung von selbst erfüllt, da alle $\lambda_\nu^{(n)}$ auf der reellen Achse liegen. Im Hermiteschen Falle kann übrigens, wie Carleman bemerkt und seinen Untersuchungen zugrunde gelegt hat, wesentlich mehr als (406) behauptet werden: Ist Ω' ein Kreis, der von der reellen Achse weder geschnitten noch berührt wird, so kann die Schranke $K(\Omega',\mathfrak{A})$ unabhängig von $\mathfrak{A}$ gewählt werden. Wir zeigen, daß entsprechendes auch ohne die Annahme des Hermiteschen Charakters gilt, so daß $\mathfrak{A}$ in (406) gestrichen werden kann. Es kann einfach

(407) $$K(\Omega',\mathfrak{A})=\frac{1}{\varepsilon-\varepsilon'}$$

gesetzt werden, wobei ε' den Radius von Ω' und ε den Radius des mit Ω' konzentrischen, Ω' enthaltenden Lückenkreises Ω bezeichnet. Zunächst ist nämlich die Matrix $\mathfrak{A}_{[n]}$ nach Voraussetzung normal, also ist es offenbar auch $\lambda\mathfrak{E}_{[n]}-\mathfrak{A}_{[n]}$ und daher auch $(\lambda\mathfrak{E}_{[n]}-\mathfrak{A}_{[n]})^{-1}=\| R_{ik}^{(n)}(\lambda)\|_{(n)}$, so daß $\mathsf{M}(\| R_{ik}^{(n)}(\lambda)\|_{(n)})$ der Toeplitzschen Polygonregel zufolge einfach den Betrag des absolut größten Eigenwertes von $\| R_{ik}^{(n)}(\lambda)\|$, oder, was nach § 8 dasselbe ist, den Betrag des absolut kleinsten Eigenwertes von $\| R_{ik}^{(n)}(\lambda)\|^{-1}=(\lambda\mathfrak{E}_{[n]}-\mathfrak{A}_{[n]})$ bezeichnet. (406) und (407) besagen also nur, daß die Eigenwerte von $(\lambda\mathfrak{E}_{[n]}-\mathfrak{A}_{[n]})$ absolut $>\varepsilon-\varepsilon'$ sind, sobald sich λ in Ω' befindet. Die Eigenwerte von $(\lambda\mathfrak{E}_{[n]}-\mathfrak{A}_{[n]})$ sind aber, wegen

$$\det[(\lambda-\lambda_\nu^{(n)})\mathfrak{E}_{[n]}-(\lambda\mathfrak{E}_{[n]}-\mathfrak{A}_{[n]})]=(-1)^n\det(\lambda_\nu^{(n)}\mathfrak{E}_{[n]}-\mathfrak{A}_{[n]})$$

und da die $\lambda_\nu^{(n)}$ die Eigenwerte von $\mathfrak{A}_{[n]}$ sind, ersichtlich die Zahlen $\lambda-\lambda_\nu^{(n)}$, sie sind also absolut $\geqq \big|\,|\lambda-\lambda_*|-|\lambda_*-\lambda_\nu^{(n)}|\,\big|>|\varepsilon'-\varepsilon|=\varepsilon-\varepsilon'$. Denn $\lambda_\nu^{(n)}$ liegt nicht im Lückenkreise Ω, d. h. es gilt $|\lambda_*-\lambda_\nu^{(n)}|\geqq\varepsilon$,

und andererseits liegt λ nach Voraussetzung auf Ω', d. h. im Kreise $|\lambda - \lambda_*| < \varepsilon' (< \varepsilon)$. — Damit ist alles bewiesen.

Von jetzt an betrachten wir nur Hermitesche Matrizen. Dies soll nicht immer wiederholt werden.

§ 103. Anwendung auf die Hermiteschen Matrizen. Grenzresolvente.

Ist $\mathfrak{A}$ Hermitesch, also $\lambda_\nu^{(n)}$ reell, so folgt aus (406), (407)

(408) $$\mathsf{M}(\| R_{ik}^{(n)}(\lambda) \|_{(n)}) \leqq \frac{1}{\mathfrak{J}[\lambda]},$$

also offenbar auch

(409) $$\mathsf{M}(\| \lim R_{ik}^{(h_n)}(\lambda) \|) \leqq \frac{1}{\mathfrak{J}[\lambda]} \qquad (\lim = \lim_{n=\infty}),$$

wobei $\mathfrak{J}[\lambda]$ den Imaginärteil von λ bezeichnet.

Hierbei ist die Grenzresolvente $\| \lim R_{ik}^{(h_n)}(\lambda) \|$ für alle nicht reellen λ bei passendem $\{h_n\}$ vorhanden und beschränkt. Insbesondere hat nach (405) jede, der Q-Bedingung genügende Hermitesche Matrix eine Grenzsolvente, die dabei eine vordere und zugleich auch hintere beschränkte Reziproke von $\lambda\mathfrak{E} - \mathfrak{A}$ darstellt und zugleich für $(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{c}$, $|\mathfrak{c}|^2 < +\infty$ eine quadratisch konvergente Lösung liefert. Solange $\mathfrak{A}$ beschränkt war, konnten wir aus den Formalsätzen oder eigentlich aus den Faltungssätzen schließen (S. 171), daß das Auswahlverfahren überflüssig ist, indem $h_n = n$ gesetzt werden kann, und daß es weiter außer der durch die Abschnitte gelieferten Grenzsolvente $\| \lim R_{ik}^{(n)}(\lambda) \|$ keine weitere beschränkte Resolvente gibt.

Wir denken dabei immer an nicht reelle λ und sagen, die Hermitesche Q-Matrix $\mathfrak{A}$ lasse die Matrix $\| R_{ik}(\lambda) \|$ als eine beschränkte Resolvente zu, wenn die folgenden Bedingungen gleichzeitig erfüllt sind: Jedes $R_{ik}(\lambda)$ ist (für nicht reelle λ) regulär, die Matrix $\| R_{ik}(\lambda) \|$ ist bei jedem (nicht reellen) λ beschränkt, es gilt $(\lambda\mathfrak{E} - \mathfrak{A}) \| R_{ik}(\lambda) \| = \mathfrak{E}$ sowie $\| R_{ik}(\lambda) \| (\lambda\mathfrak{E} - \mathfrak{A}) = \mathfrak{E}$ [wobei die Produkte, da beide Faktoren Q-Matrizen sind, nach S. 122 eindeutig gebildet werden können], und $x_i = \sum_{k=1}^{\infty} R_{ik}(\lambda) c_k$ stellt, sobald $\sum_{i=1}^{\infty} |c_i|^2 < +\infty$ ist, eine (offenbar quadratisch konvergente) Lösung von $(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{c}$ dar [wobei der Vektor $(\lambda\mathfrak{E} - \mathfrak{A})\mathfrak{x}$, der Schwarzschen Ungleichheit zufolge und weil $\lambda\mathfrak{E} - \mathfrak{A}$ eine Q-Matrix ist, wieder eindeutig gebildet werden kann].

Bei nicht beschränkten Hermiteschen Q-Matrizen $\mathfrak{A}$ ist nun die Sache nicht so einfach wie bei der beschränkten. Es sind jetzt nämlich, wie Beispiele zeigen (vgl. unten), alle drei denkbaren Fälle möglich (je nach der Wahl von $\mathfrak{A}$): a) Möglicherweise gibt es ebenfalls nur eine beschränkte Resolvente, also erst recht nur eine Grenzresolvente, so daß das Auswahlverfahren gewiß überflüssig ist. — b) Es ist aber auch möglich, daß es mehrere beschränkte Resolventen gibt, trotzdem das Auswahlverfahren überflüssig ist, so daß beschränkte Resolventen existieren, die von der beschränkten Grenzresolvente $\|\lim R_{ik}^{(n)}(\lambda)\|$ verschieden sind, so daß also ihre Existenz nicht bereits durch die Betrachtung der Abschnitte $\mathfrak{A}_{[n]}$ festgestellt werden kann. — c) Endlich ist es sehr wohl möglich, daß die Existenz mehrerer beschränkter Resolventen bereits durch das Abschnittsprinzip angezeigt wird, weil sogar *mehrere* (eo ipso beschränkte) *Grenz*resolventen $\|\lim R_{ik}^{(h_n)}(\lambda)\|$ vorhanden sind, so daß nicht alle Auswahlfolgen $\{h_n\}$ zu derselben Grenzresolvente führen; m. a. W., man kann nicht $h_n = n$ setzen, da $\|\lim R_{ik}^{(n)}(\lambda)\|$ nicht existiert.

§ 104. Reguläre Matrizen.

Eine Hermitesche Q-Matrix, für welche der Fall a) vorliegt, nennen wir *regulär*. Z. B. ist jede beschränkte Hermitesche Matrix regulär. Daß es auch nicht reguläre Matrizen gibt, ist nicht erstaunlich, der Leser ist darauf in Kap. III mehrfach vorbereitet worden. Alle scheinbaren Widersprüche bei nicht beschränkten Q-Matrizen erklären sich, wenn man einfach das Rechnen mit divergenten Doppelreihen bzw. nicht begründete Umordnungen von mehrfachen Reihen vermeidet und gliedweise Vornahme von Grenzübergängen im Falle einer evtl. nur schwachen Konvergenz unterläßt. Der Sachverhalt ist der folgende. Die Matrix $\mathfrak{A}$ ist dann und nur dann beschränkt, wenn etwa die zweifache Reihe $\sum_{i=1}^{\infty}(\sum_{k=1}^{\infty} a_{ik} x_i \bar{x}_k)$ in jedem Punkte von E konvergiert (und außerdem evtl. auch für gewisse, nicht quadratisch konvergente Argumente $\{x_i\}$); und zwar ist dann die Reihe $\sum_{k=1}^{\infty}(\sum_{i=1}^{\infty} a_{ik} x_i \bar{x}_k)$ auf E ebenfalls konvergent und gleich der vorigen; vgl. S. 125. Ist also $\mathfrak{A}$ nicht beschränkt, so gibt es Punkte auf E, wo keine dieser beiden Reihen konvergiert. Ist $\mathfrak{A}$ beschränkt, so ist ferner $T = \sum_{i=1}^{\infty} |\sum_{k=1}^{\infty} a_{ik} \bar{x}_k|^2$ auf E konvergent. Ist aber $\mathfrak{A}$ nicht beschränkt, so gibt es Punkte auf E,

wo T divergiert. Es bezeichne $\mathsf{E}_{\mathfrak{A}}$ diejenige Teilmenge von E, wo T konvergiert. $\mathsf{E}_{\mathfrak{A}}$ ist gewiß nicht leer, da z. B. jedes $\mathsf{E}_{[n]}$ auf $\mathsf{E}_{\mathfrak{A}}$ liegt. Auf $\mathsf{E}_{\mathfrak{A}}$ ist der Schwarzschen Ungleichheit zufolge die erste der beiden Reihen

$$(410) \qquad \sum_{i=1}^{\infty}\Big(\sum_{k=1}^{\infty} a_{ik}\, x_i\, \bar{x}_k\Big), \qquad \sum_{k=1}^{\infty}\Big(\sum_{i=1}^{\infty} a_{ik}\, x_i\, \bar{x}_k\Big)$$

gewiß konvergent. Eine jede der inneren Summen ist, da $\mathfrak{A} = \mathfrak{A}^*$ eine Q-Matrix, also $\sum_{k=1}^{\infty} |a_{ik}|^2 = \sum_{k=1}^{\infty} |a_{ki}|^2 < +\infty$ ist, der Schwarzschen Ungleichheit zufolge nicht nur auf $\mathsf{E}_{\mathfrak{A}}$, sondern auch auf E konvergent, und zwar absolut.

Es sind nun drei Fälle denkbar: I) Beide Reihen (410) sind auf $\mathsf{E}_{\mathfrak{A}}$ konvergent und liefern auf $\mathsf{E}_{\mathfrak{A}}$ denselben Wert für die „Kopplungsform". Es ist ferner denkbar, daß II) auf $\mathsf{E}_{\mathfrak{A}}$ zwar beide Reihen (410) konvergent sind, daß es aber Punkte auf $\mathsf{E}_{\mathfrak{A}}$ gibt, wo die beiden Reihen verschiedene Werte ergeben, oder endlich III), daß es Punkte auf $\mathsf{E}_{\mathfrak{A}}$ gibt, wo nur eine der Reihen (410) konvergiert, die andere aber divergiert. Wir zeigen nun mit Carleman, daß die Annahme I) mit der Bedingung a), d. h. mit der Regularitätsvoraussetzung äquivalent ist.

Die Hermitesche Q-Matrix $\mathfrak{A}$ ist nach dem auf S. 221 Gesagten dann und nur dann regulär, wenn das System $(\lambda\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{x} = \mathfrak{c}$ bei jedem quadratisch konvergenten $\mathfrak{c}$ und bei jedem nicht reellen λ nur *eine* quadratisch konvergente Lösung $\mathfrak{x}$ hat. Ist also $\mathfrak{A}$ regulär, so lassen die homogenen Gleichungen für nicht reelle λ nur die triviale Lösung $\mathfrak{x} = \mathfrak{o}$ zu; denn der Nullvektor ist *eine* quadratisch konvergente Lösung, also nach Voraussetzung die *einzige*. Ist umgekehrt $\mathfrak{A}$ nicht regulär, so gibt es ein nicht reelles $\lambda = \lambda_0$ und dazu ein $\mathfrak{c}$ derart, daß $(\lambda_0\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{x} = \mathfrak{c}$ sowohl durch $\mathfrak{x} = \mathfrak{x}_1$ als auch durch $\mathfrak{x} = \mathfrak{x}_2$ erfüllt werden kann, wobei $|\mathfrak{x}_1|^2 < +\infty$, $|\mathfrak{x}_2|^2 < +\infty$, $\mathfrak{x}_1 \neq \mathfrak{x}_2$ gilt, so daß $\mathfrak{x}_0 = \mathfrak{x}_1 - \mathfrak{x}_2$ offenbar eine quadratisch konvergente, von $\mathfrak{o}$ verschiedene Lösung der homogenen Gleichungen darstellt. Es wird sich übrigens herausstellen, daß hierbei λ_0 beliebig ist, so daß, wenn $\mathfrak{A}$ nicht regulär ist, die homogenen Gleichungen $(\lambda_0\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{x} = \mathfrak{o}$ bei *jedem* nicht reellen λ_0 Lösungen $\mathfrak{x}$ mit $0 < |\mathfrak{x}|^2 < +\infty$ zulassen müssen, m. a. W., die Gleichungen $(\lambda_0\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{x} = \mathfrak{c}$ haben dann bei jedem nicht reellen λ_0 und bei jedem quadratisch konvergenten $\mathfrak{c}$ mehrere quadratisch konvergente Lösungen.

§ 105. Fortsetzung. Beweis des Carlemanschen Kriteriums.

Es sei also $\mathsf{E}_{\mathfrak{A}}$ die oben erklärte Teilmenge von E, auf welcher die Reihe $\sum_{i=1}^{\infty}(\sum_{k=1}^{\infty} a_{ik} x_i \bar{x}_k)$ mit Rücksicht auf die Schwarzsche Ungleichung offenbar konvergiert. Wir haben zu zeigen, daß die zu $\lambda_0 \mathfrak{E} - \mathfrak{A}$; $\mathfrak{J}[\lambda_0] \neq 0$ gehörigen homogenen Gleichungen dann und nur dann eine nicht triviale $(\neq \mathfrak{o})$ quadratisch konvergente Lösung zulassen, wenn auf $\mathsf{E}_{\mathfrak{A}}$ auch die Reihe $\sum_{k=1}^{\infty}(\sum_{i=1}^{\infty} a_{ik} x_i \bar{x}_k)$ konvergiert und denselben Wert wie die vorige liefert (die Bedingung ist ersichtlich unabhängig von λ_0).

Um zunächst den zweiten Teil des Satzes zu beweisen, nehmen wir an, daß auf $\mathsf{E}_{\mathfrak{A}}$ durchweg

$$\sum_{i=1}^{\infty}(\sum_{k=1}^{\infty} a_{ik} x_i \bar{x}_k) = \sum_{k=1}^{\infty}(\sum_{i=1}^{\infty} a_{ik} x_i \bar{x}_k) \tag{411}$$

gilt und setzen zwecks Zurückführung ad absurdum voraus, daß die homogenen Gleichungen

$$\lambda \bar{\xi}_i - \sum_{k=1}^{\infty} a_{ik} \bar{\xi}_k = 0 \tag{412}$$

bei irgendeinem nicht reellen λ eine von $\mathfrak{o}$ verschiedene quadratisch konvergente Lösung, also auch eine Lösung mit $\sum_{i=1}^{\infty} |\xi_i|^2 = 1$ haben. Aus (412) folgt, daß die Reihe

$$\sum_{i=1}^{\infty}(\sum_{k=1}^{\infty} a_{ik} \xi_i \bar{\xi}_k) = \lambda \sum_{i=1}^{\infty} \xi_i \bar{\xi}_i = \lambda \tag{413}$$

konvergiert. Aus (412) folgt aber auch

$$T = \sum_{i=1}^{\infty} |\sum_{k=1}^{\infty} a_{ik} \bar{\xi}_k|^2 = |\lambda|^2,$$

so daß der Punkt $\{\xi_i\}$ nach Definition auf $\mathsf{E}_{\mathfrak{A}}$ gelegen ist. Geht man in (412) zu den konjugierten Größen über, so folgt

$$\bar{\lambda} \xi_i - \sum_{k=1}^{\infty} \bar{a}_{ik} \xi_k = 0,$$

wofür wegen $a_{ik} = \bar{a}_{ki}$ auch

$$\bar{\lambda} \xi_k - \sum_{i=1}^{\infty} a_{ik} \xi_i = 0 \tag{414}$$

geschrieben werden kann. Aus (414) folgt nun

$$(415)\qquad \sum_{k=1}^{\infty}\Big(\sum_{i=1}^{\infty} a_{ik}\,\xi_i\,\bar{\xi}_k\Big) = \bar{\lambda}\sum_{k=1}^{\infty}\xi_i\,\bar{\xi}_k = \bar{\lambda}\,.$$

Nun liegt aber $\{\xi_i\}$ auf $\mathsf{E}_{\mathfrak{A}}$, wegen (411), (413), (415) muß also $\lambda = \bar{\lambda}$, d. h. λ reell sein, im Widerspruch mit der Voraussetzung.

Um auch die Umkehrung zu beweisen, nehmen wir an, daß es ein nicht reelles λ gibt derart, daß die zu $\lambda\,\mathfrak{E} - \mathfrak{A}$ gehörigen homogenen Gleichungen keine von dem Nullvektor verschiedene Lösung und daher die inhomogenen Gleichungen nur eine Lösung haben, unter Lösung hier und später stets nur eine quadratisch konvergente Lösung verstanden. Wir haben zu beweisen, daß dann auf $\mathsf{E}_{\mathfrak{A}}$ der Vertauschungssatz (411), d. h. $\mathfrak{x}\cdot\mathfrak{A}\,\bar{\mathfrak{x}} = \mathfrak{x}\cdot\mathfrak{A}'\,\bar{\mathfrak{x}}$ gilt, oder, was wegen $\mathfrak{A} = \mathfrak{A}^* = \bar{\mathfrak{A}}'$ offenbar dasselbe ist, daß $\mathfrak{x}\cdot\mathfrak{A}\,\bar{\mathfrak{x}} = \bar{\mathfrak{x}}\cdot\mathfrak{A}\,\mathfrak{x}$ ausfällt. Dann wird bewiesen sein, daß diese Symmetriebedingung auf $\mathsf{E}_{\mathfrak{A}}$, wie behauptet, nicht nur hinreichend, sondern auch notwendig ist für die Regularität der Hermiteschen Q-Matrix $\mathfrak{A}$. Wir beschränken uns dabei auf den Fall, wo die Hermitesche Matrix $\mathfrak{A}$, und daher auch jede Abschnittsresolvente, mithin auch die Grenzresolvente mit ihrer transponierten Matrix identisch ist.

Es seien $\mathfrak{x}$, $\mathfrak{y}$ irgend zwei Vektoren auf $\mathsf{E}_{\mathfrak{A}}$. Letztere Mannigfaltigkeit ist, wie erinnerlich, dadurch definiert (S. 222), daß sie diejenigen und nur diejenigen Vektoren von E enthält, welche durch $\mathfrak{A}$ in einen quadratisch konvergenten Vektor $\mathfrak{A}\,\mathfrak{x}$ übergeführt werden. Denn $\dot{T} < +\infty$ ist gleichbedeutend mit $|\,\mathfrak{A}\,\mathfrak{x}\,|^2 < +\infty$. Es sind daher die Vektoren $\mathfrak{A}\,\mathfrak{x}$, $\mathfrak{A}\,\mathfrak{y}$ gewiß quadratisch konvergent. Aber auch die Vektoren $\mathfrak{x}$, $\mathfrak{y}$ sind quadratisch konvergent (denn sie liegen auf E), mithin sind es auch die beiden Vektoren $\mathfrak{c} = \lambda\,\mathfrak{x} - \mathfrak{A}\,\mathfrak{x}$, $\mathfrak{d} = \lambda\,\mathfrak{y} - \mathfrak{A}\,\mathfrak{y}$, wobei λ jene vorgegebene nicht reelle Zahl ist. Offenbar gilt

$$(416)\qquad (\lambda\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{x} = \mathfrak{c}\,,\qquad (\lambda\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{y} = \mathfrak{d}\,,$$

also, da die Lösung nach Voraussetzung eindeutig ist,

$$(417)\qquad \mathfrak{x} = \mathfrak{R}_\lambda\,\mathfrak{c}\,,\qquad \mathfrak{y} = \mathfrak{R}_\lambda\,\mathfrak{d}\,,$$

woselbst $\mathfrak{R}_\lambda = (\lambda\,\mathfrak{E} - \mathfrak{A})^{-1}$ die Grenzresolvente von $\mathfrak{A}$ bezeichnet (vgl. S. 220), also beschränkt ist. Die zu $\mathfrak{R}_\lambda$ gehörige Bilinearform kann daher nach (216), S. 125 für die quadratisch konvergenten Argumente $\mathfrak{c}$, $\mathfrak{d}$ sowohl zeilenweise als auch kolonnenweise summiert werden, d. h. man hat $\mathfrak{d}\cdot\mathfrak{R}_\lambda\,\mathfrak{c} = \mathfrak{c}\cdot\mathfrak{R}_\lambda\,\mathfrak{d}$, also wegen (416) und (417) offenbar $[(\lambda\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{y}]\cdot\mathfrak{x} = [(\lambda\,\mathfrak{E} - \mathfrak{A})\,\mathfrak{x}]\cdot\mathfrak{y}$, wobei der Punkt skalare Vektormultiplikation bezeichnet. Wegen $(\mathfrak{C} + \mathfrak{D})\,\mathfrak{a} = \mathfrak{C}\,\mathfrak{a} + \mathfrak{D}\,\mathfrak{a}$ und $\mathfrak{a}\cdot\mathfrak{b} = \mathfrak{b}\cdot\mathfrak{a}$

ergibt dies nach Subtraktion von $(\lambda \mathfrak{E} \mathfrak{y}) \cdot \mathfrak{x} = \lambda \mathfrak{y} \cdot \mathfrak{x} = \lambda \mathfrak{x} \cdot \mathfrak{y} = (\lambda \mathfrak{E} \mathfrak{x}) \cdot \mathfrak{y}$, wie leicht ersichtlich, $(- \mathfrak{A} \mathfrak{y}) \cdot \mathfrak{x} = (- \mathfrak{A} \mathfrak{x}) \cdot \mathfrak{y}$, d. h. $\mathfrak{x} \cdot \mathfrak{A} \mathfrak{y} = \mathfrak{y} \cdot \mathfrak{A} \mathfrak{x}$. Dies gilt für zwei beliebige Vektoren $\mathfrak{x}$, $\mathfrak{y}$ in $\mathsf{E}_{\mathfrak{A}}$. Nun liegt aber, des Hermiteschen Charakters zufolge, zugleich mit $\mathfrak{x}$ auch $\bar{\mathfrak{x}}$ auf $\mathsf{E}_{\mathfrak{A}}$, so daß wir in der letzten Beziehung $\mathfrak{y} = \bar{\mathfrak{x}}$ setzen dürfen, w. z. b. w.

Durch eine sinngemäße Wiederholung dieser Schlußweise ließe sich zeigen, daß $\mathfrak{R}_\lambda$ bei regulären Matrizen gewiß der Hilbertschen Funktionalgleichung genügt und man könnte dann auch die in § 79 durchgeführte Rechnung wiederholen.

§ 106. Statistisch sinnvolle Matrizen. Grenzspektralmatrizen.

Eine Hermitesche, sonst beliebige Matrix $\mathfrak{A}$, für welche *nicht* der Fall c) [sondern entweder a) oder b)] vorliegt, bei welcher also der Grenzwert

$$(418) \qquad \lim_{n=\infty} R_{ik}^{(n)}(\lambda) = \lim_{n=\infty} \int_{-\infty}^{+\infty} \frac{d\sigma_{ik}^{(n)}(\mu)}{\lambda - \mu},$$

wobei $\|\sigma_{ik}^{(n)}\|_{(n)}$ die Spektralmatrix von $\mathfrak{A}_{[n]}$ bezeichnet, für alle λ ($\mathfrak{J}[\lambda] \neq 0$) existiert, wollen wir *statistisch sinnvoll* nennen (auch wenn die Q-Bedingung nicht erfüllt ist). Es ist nämlich $|\sigma_{ik}^{(n)}(\mu)| \leqq 1$ und $\int_{-\infty}^{+\infty} |d\sigma_{ik}^{(n)}(\mu)| \leqq 1$ (siehe S. 51), so daß mit Rücksicht auf das Fundamentalkriterium (§ 48) $\lim R_{ik}^{(h_m)}(\lambda)$ für alle nicht reellen λ dann und nur dann existiert, wenn $\lim \sigma_{ik}^{(h_m)}(\mu)$ vorhanden ist $(m \to +\infty)$; von den Unstetigkeitsstellen der Grenzfunktion und von einer additiven Konstante hat man dabei möglicherweise abzusehen. Ist ferner das Auswahlverfahren für die $R_{ik}^{(n)}$ nicht überflüssig, so ist es auch für die $\sigma_{ik}^{(n)}$ nicht, und umgekehrt. Die Matrix $\mathfrak{A}$ ist also dann und nur dann statistisch sinnvoll, wenn das durch $\|\sigma_{ik}^{(n)}\|_{(n)}$ beschriebene statistische Verteilungsgesetz für $n \to +\infty$ einer *bestimmten* Grenzlage zustrebt, während sonst eine *Mannigfaltigkeit* von Grenzstatistiken vorliegt. Daher der Name: statistisch sinnvoll.

Ist $\mathfrak{A}$ statistisch sinnvoll, so gibt es eine Funktion $\varrho_{ik}(\mu)$ derart, daß im Sinne der Hellyschen wesentlichen Konvergenz $\sigma_{ik}^{(n)} [\to] \varrho_{ik}$ gilt. Und zwar ist dabei $|\varrho_{ik}| \leqq 1$, $\int_{-\infty}^{+\infty} |d\varrho_{ik}(\mu)| \leqq 1$. Wir legen ϱ_{ik} an den Unstetigkeitsstellen wieder durch die Forderung der rechts-

seitigen Stetigkeit fest, bringen den wegen $\int\limits_{-\infty}^{+\infty} |d\varrho_{ik}| < +\infty$ gewiß vorhandenen Grenzwert $\varrho_{ik}(-\infty)$ in Abzug und setzen

$$\overset{*}{\sigma}_{ik}(\mu) = \varrho_{ik}(\mu + 0) - \varrho_{ik}(-\infty). \tag{419}$$

Die dadurch festgelegte Matrix $\| \overset{*}{\sigma}_{ik}(\mu) \|$ von rechtsseitig stetigen, im negativen Unendlichen verschwindenden Funktionen nennen wir die Grenzspektralmatrix von $\mathfrak{A}$. Bei statistisch sinnvollen (und übrigens auch nur bei solchen) Matrizen gibt es also nur eine Grenzspektralmatrix. Offenbar ist $\mathsf{M}(\| \overset{*}{\sigma}_{ik}(\mu) \|)$ eine nicht abnehmende Funktion von μ; genauer: $\sum\limits_{i=1}^{\infty} \sum\limits_{k=1}^{\infty} \overset{*}{\sigma}_{ik}(\mu) x_i \bar{x}_k = \Phi(\| \overset{*}{\sigma}_{ik}(\mu) \|; \mathfrak{x})$ ist, wenn $\mathfrak{x}$ einen festen Punkt auf E bezeichnet, eine ebenfalls nicht abnehmende Funktion von μ, die durchweg $\leqq 1$ und $\geqq 0$ ist. Dies alles folgt unmittelbar daraus, daß entsprechendes für die $\| \sigma_{ik}^{(n)}(\mu) \|_{(n)}$, wobei $\| \sigma_{ik}^{(n)}(-\infty) \|_{(n)} = \| 0 \|_{(n)}$ ist, gilt. Insbesondere ist also $\| \overset{*}{\sigma}_{ik}(\mu) \|$ auch dann beschränkt, wenn $\mathfrak{A}$ nicht beschränkt ist.

Für nicht reelle λ gilt dabei für die Grenzresolvente

$$\| \lim R_{ik}^{(n)}(\lambda) \| = \| R_{ik}(\lambda) \| \tag{420}$$

die Darstellung

$$R_{ik}(\lambda) = \int\limits_{-\infty}^{+\infty} \frac{d\overset{*}{\sigma}(\mu)}{\lambda - \mu}. \tag{421}$$

Der Beweis verläuft, unter Benutzung des Fundamentalkriteriums, wörtlich so wie auf S. 169 ff.

§ 107. Integraldarstellung der Q-Matrizen.

Ist $\mathfrak{A} = \| a_{ik} \|$ eine statistisch sinnvolle Q-Matrix mit der Grenzspektralmatrix $\| \overset{*}{\sigma}_{ik}(\mu) \|$, so gilt

$$\delta_{ik} = \int\limits_{-\infty}^{+\infty} d\overset{*}{\sigma}_{ik}(\mu), \qquad a_{ik} = \int\limits_{-\infty}^{+\infty} \mu \, d\overset{*}{\sigma}_{ik}(\mu). \tag{422}$$

Hingegen ist eventuell bereits das Integral $\int\limits_{-\infty}^{+\infty} \mu^3 \, d\sigma_{ik}(\mu)$ divergent, so daß (305) nicht für alle ν zu gelten braucht. Dem entspricht es, daß $\mathfrak{A}^2$ zwar existiert, aber vielleicht schon keine Q-Matrix ist, so daß

$\mathfrak{A}^2\mathfrak{A} = \mathfrak{A}^3$ vielleicht überhaupt nicht gebildet werden kann. Wie wir später sehen werden, hängen die Verhältnisse mit einem auf S. 102 erwähnten Umstand zusammen.

Wir gehen zum Beweis der Formeln (422) über, die von Carleman herrühren. Es bezeichne $\| \sigma_{ik}^{(n)}(\mu) \|_{(n)}$ die Spektralmatrix von $\mathfrak{A}_{[n]}$, so daß

$$(64)' \qquad \sigma_{ik}^{(n)}(\mu) = \sum_{\lambda_\nu^{(n)} \leqq \mu} u_{i\nu}^{(n)} \overline{u}_{k\nu}^{(n)}$$

gilt, wobei die $\lambda_\nu^{(n)}$ die Abschnittseigenwerte bezeichnen und die $u_{i\nu}^{(n)}$ zu $\lambda_\nu^{(n)} \mathfrak{E}_{[n]} - \mathfrak{A}_{[n]}$ gehörige normiert orthogonale Eigenlösungen bedeuten; vgl. S. 32. Es sei $\gamma > 0$ beliebig groß und es bezeichne α, β ein Zahlenpaar derart, daß entweder $\gamma < \alpha < \beta$ oder $\alpha < \beta < -\gamma$ gilt. Dann ist jeder der Bedingung $\alpha \leqq \lambda_l^{(n)} \leqq \beta$ genügende Eigenwert absolut $> \gamma$, so daß — wenn man die Zeiger der Bequemlichkeit halber auf eine leicht ersichtliche Weise vertauscht — mit der Abkürzung $C_i^2 = \sum_{j=1}^{\infty} |a_{ij}|^2$, $C_i \geqq 0$ offenbar

$$\Big| \sum_{\alpha \leqq \lambda_l^{(n)} \leqq \beta} u_{il}^{(n)} \overline{u}_{kl}^{(n)} \Big|^2$$

$$= \Big| \sum_{\alpha \leqq \lambda_l^{(n)} \leqq \beta} \overline{u}_{kl}^{(n)} \frac{1}{\lambda_l^{(n)}} \sum_{j=1}^{n} a_{ij} u_{jl}^{(n)} \Big|^2 = \Big| \sum_{j=1}^{n} a_{ij} \sum_{\alpha \leqq \lambda_l^{(n)} \leqq \beta} \frac{u_{jl}^{(n)} \overline{u}_{kl}^{(n)}}{\lambda_l^{(n)}} \Big|^2$$

$$\leqq \Big(\sum_{j=1}^{n} |a_{ij}|^2 \Big) \sum_{j=1}^{n} \Big| \sum_{\alpha \leqq \lambda_l^{(n)} \leqq \beta} \frac{u_{jl}^{(n)} \overline{u}_{kl}^{(n)}}{\lambda_l^{(n)}} \Big|^2$$

$$\leqq C_i^2 \sum_{j=1}^{n} \sum_{\alpha \leqq \lambda_l^{(n)} \leqq \beta} \frac{u_{jl}^{(n)} \overline{u}_{kl}^{(n)}}{\lambda_l^{(n)}} \sum_{\alpha \leqq \lambda_m^{(n)} \leqq \beta} \frac{\overline{u}_{jm}^{(n)} u_{km}^{(n)}}{\lambda_m^{(n)}}$$

$$= C_i^2 \sum_{\alpha \leqq \lambda_l^{(n)} \leqq \beta} \sum_{\alpha \leqq \lambda_m^{(n)} \leqq \beta} \frac{u_{km}^{(n)} \overline{u}_{kl}^{(n)}}{\lambda_m^{(n)} \lambda_l^{(n)}} \sum_{j=1}^{n} u_{jl}^{(n)} \overline{u}_{jm}^{(n)}$$

$$= C_i^2 \sum_{\alpha \leqq \lambda_l^{(n)} \leqq \beta} \sum_{\alpha \leqq \lambda_m^{(n)} \leqq \beta} \frac{u_{km}^{(n)} \overline{u}_{kl}^{(n)}}{\lambda_m^{(n)} \lambda_l^{(n)}} \delta_{lm} = C_i^2 \sum_{\alpha \leqq \lambda_l^{(n)} \leqq \beta} \frac{u_{kl}^{(n)} \overline{u}_{kl}^{(n)}}{\lambda_l^{(n)} \lambda_l^{(n)}}$$

$$\leqq \frac{C_i^2}{\gamma^2} \sum_{\alpha \leqq \lambda_l^{(n)} \leqq \beta} |u_{kl}^{(n)}|^2 \leqq \frac{C_i^2}{\gamma^2} \sum_{l=1}^{n} |u_{kl}^{(n)}|^2 = \frac{C_i^2}{\gamma^2},$$

d. h. wegen (64)'

$$\Big| \underset{[\alpha, \beta]}{\Delta} \sigma_{ik}^{(n)} \Big| \leqq \frac{C_i}{\gamma},$$

mithin

$$(*)\qquad \left|\int_{\gamma}^{+\infty} d\sigma_{ik}^{(n)}(\mu)\right| + \left|\int_{-\infty}^{-\gamma} d\sigma_{ik}^{(n)}(\mu)\right| < \frac{\varepsilon}{3}$$

gilt, wenn γ hinreichend groß ist; ε bedeutet dabei eine beliebig kleine, von n unabhängige Zahl, die evtl. von i abhängt. Es sei γ der Einfachheit halber immer derart, daß $\mu = \pm\gamma$ keine Sprungstelle der $\sigma_{ik}^{(n)}$ oder von $\overset{*}{\sigma}_{ik}$ ist (solche γ gibt es überall dicht, da es ja insgesamt nur abzählbar viele Sprungstellen gibt). Für hinreichend große γ gilt wegen $\int_{-\infty}^{+\infty} |d\sigma_{ik}^{(n)}(\mu)| \leqq 1$

$$(*)\qquad \left|\int_{\gamma}^{+\infty} d\overset{*}{\sigma}_{ik}(\mu)\right| + \left|\int_{-\infty}^{-\gamma} d\overset{*}{\sigma}_{ik}(\mu)\right| < \frac{\varepsilon}{3}.$$

Ferner kann man, nach geschehener Wahl von γ, mit Rücksicht auf den Hellyschen Satz über gliedweise Integration, ein N derart festlegen, daß für $n > N$

$$(*)\qquad \left|\int_{-\gamma}^{+\gamma} d\sigma_{ik}^{(n)}(\mu) - \int_{-\gamma}^{+\gamma} d\overset{*}{\sigma}_{ik}(\mu)\right| < \frac{\varepsilon}{3}$$

ausfällt. Aus den drei, mit (*) bezeichneten Ungleichungen folgt offenbar

$$\int_{-\infty}^{+\infty} d\sigma_{ik}^{(n)}(\mu) \to \int_{-\infty}^{+\infty} d\overset{*}{\sigma}_{ik}(\mu),$$

womit die erste Relation (422) bewiesen ist (da wir in $\int_{-\gamma_1}^{\gamma_2}$ nicht notwendig $\gamma_1 = \gamma_2$ vorauszusetzen brauchen). Denn in der letzten Grenzgleichung ist das Integral linkerhand nach S. 51 bei jedem n gleich δ_{ik}.

Da $\mathsf{M}(\|\sigma_{ik}^{[n]}(\mu)\|_{(n)}) \geqq 0$ eine nicht abnehmende Funktion und $\leqq 1$ ist, so schließt man wie auf S. 175, daß $\sum_{j=1}^{n} |\underset{\lambda_j^{(n)}}{\Delta} \sigma_{ik}|^2$ unterhalb einer festen Schranke bleibt, so daß mit Rücksicht auf das Majorantenkriterium der starken Konvergenz das Integral

$$\int_{-\gamma}^{\gamma} \mu\, d\sigma_{ik}^{(n)}(\mu) = \sum_{-\gamma<\lambda_l^{(n)}<\gamma} \lambda_l^{(n)} \underset{\lambda_l^{(n)}}{\Delta} \sigma_{ik}^{(n)} = \sum_{-\gamma<\lambda_l^{(n)}<\gamma} \lambda_l^{(n)} u_{il}^{(n)} \overline{u}_{kl}^{(n)}$$

$$= \sum_{-\gamma<\lambda_l^{(n)}<\gamma} \overline{u}_{kl}^{(n)} \sum_{j=1}^{n} a_{ij} u_{jl}^{(n)} = \sum_{j=1}^{n} a_{ij} \sum_{-\gamma<\lambda_l^{(n)}<\gamma} \underset{\lambda_l^{(n)}}{\Delta} \sigma_{jk}^{(n)}$$

$$= \sum_{j=1}^{n} a_{ij} \{\sigma_{jk}^{(n)}(\gamma) - \sigma_{jk}^{(n)}(-\gamma)\}$$

für $n \to +\infty$ gegen

$$\sum_{j=1}^{\infty} a_{ij} \{\overset{*}{\sigma}_{jk}(\gamma) - \overset{*}{\sigma}_{jk}(-\gamma)\} = \sum_{j=1}^{\infty} a_{ij} \int_{-\gamma}^{\gamma} d\overset{*}{\sigma}_{jk}(\mu),$$

andererseits aber nach Helly gegen $\int_{-\gamma}^{\gamma} \mu \, d\overset{*}{\sigma}_{ik}(\mu)$ konvergiert. Mithin ist

$$\int_{-\gamma_1}^{\gamma_2} \mu \, d\overset{*}{\sigma}_{ik}(\mu) = \sum_{j=1}^{\infty} a_{ij} \int_{-\gamma_1}^{\gamma_2} d\overset{*}{\sigma}_{jk}(\mu).$$

Nach dem Majorantenkriterium der starken Konvergenz, dessen Prämissen augenscheinlich erfüllt sind, können hierbei die Grenzübergänge $\gamma_2 \to +\infty$, $\gamma_1 \to -\infty$ gliedweise ausgeführt werden, und es bleibt

$$\int_{-\infty}^{+\infty} \mu \, d\overset{*}{\sigma}_{ik}(\mu) = \sum_{j=1}^{\infty} a_{ij} \int_{-\infty}^{+\infty} d\overset{*}{\sigma}_{jk}(\mu) = \sum_{j=1}^{\infty} a_{ij} \delta_{jk} = a_{ik},$$

w. z. b. w. — Ist $\mathfrak{A}$ nicht statistisch sinnvoll, so daß es mehrere Grenzspektralmatrizen gibt, so gilt (422) offenbar für eine jede Grenzspektralmatrix.

§ 108. Halbbeschränkte Matrizen.

Bevor wir weitergehen, müssen wir einen grundlegenden Satz von Toeplitz vorausschicken. — Es sei $\mathfrak{B}$ eine positiv definite, sonst beliebige Hermitesche Matrix. Es wird also nur vorausgesetzt (vgl. S. 124), daß alle Abschnittseigenwerte von $\mathfrak{B}$ oberhalb einer positiven (also von Null verschiedenen) unteren Schranke liegen, die wir mit $\mathsf{n}(\mathfrak{B})$ bezeichnet haben. Es wird hingegen nicht einmal die Q-Bedingung vorausgesetzt, sondern eben nur $\mathsf{n}(\mathfrak{B}) > 0$. Es ist also

$$(423) \quad 0 < \mathsf{n}(\mathfrak{B}) \leqq \mathsf{n}(\mathfrak{B}_{[n]}) \leqq {}_0\lambda_\nu^{(n)} \leqq \mathsf{m}(\mathfrak{B}_{[n]}) \leqq \mathsf{m}(\mathfrak{B}) \leqq +\infty,$$

wobei die ${}_0\lambda_\nu^{(n)}$; $\nu = 1, 2, \ldots, n$ die Eigenwerte von $\mathfrak{B}_{[n]}$ bezeichnen. Da $\det \mathfrak{B}_{[n]}$ das Produkt der Eigenwerte, also $\neq 0$ ist, so ist die Reziproke $(\| b_{ik} \|_{[n]})^{-1} = (\mathfrak{B}_{[n]})^{-1}$, die wir kurz mit $\mathfrak{B}_{[n]}^{-1}$ bezeichnen und $= \| B_{ik}^{[n]} \|_{(n)}$ setzen wollen, vorhanden; $1 \leqq i, k \leqq n$. Der Toeplitzsche Satz besagt nun, daß der Grenzwert $\lim_{n=\infty} B_{ik}^{[n]} = B_{ik}$ bei jedem i und k existiert. Wir wissen von vornherein, daß, *wenn* dieser Grenzwert existiert, $\| B_{ik} \|$ eine beschränkte Matrix sein muß (vgl. S. 218), doch wird sich dies später von selbst ergeben. — Auf diese Weise hat Toeplitz bei beschränkten Matrizen seinen Existenzsatz auf S. 136 bewiesen, für den

wir die Hilbsche Beweisanordnung zugrunde gelegt haben. — Es ist unnötig zu betonen, daß die beiden Matrizen $\| B_{ik}^{[n]} \|_{(n)}$, $\| B_{ik} \|_{[n]}$ nicht zu verwechseln sind. Dies ist eben der springende Punkt. (Wären die beiden Matrizen identisch, so wäre der Satz eine Trivialität.) — Den Beweis des Toeplitzschen Satzes wollen wir noch verschieben und zunächst aus ihm eine Folgerung ziehen, die eine brauchbare hinreichende Bedingung liefert dafür, daß eine Hermitesche Matrix $\mathfrak{A}$ statistisch sinnvoll ist.

Es sei die Matrix $\mathfrak{A}$ Hermitesch und es seien $\lambda_\nu^{(n)}$; $\nu = 1, 2, \ldots$, $n = 1, 2, \ldots$ die Abschnittseigenwerte. *Beschränkt* ist $\mathfrak{A}$, wie wir wissen, dann und nur dann, wenn

$$-\infty < \mathsf{n}(\mathfrak{A}) \leqq \mathsf{m}(\mathfrak{A}) < +\infty$$

gilt, oder, was dasselbe ist, wenn das Abschnittsspektrum, d. h. die Gesamtheit der $\lambda_\nu^{(n)}$, durch einen *endlichen* Intervall $\mathsf{n}(\mathfrak{A}) \leqq \lambda \leqq \mathsf{m}(\mathfrak{A})$ überdeckt werden kann. Dementsprechend wollen wir die Matrix $\mathfrak{A}$ *halbbeschränkt* nennen, wenn das Abschnittsspektrum von einer passenden *Halbgeraden* überdeckt wird, so daß entweder

$$(424) \qquad -\infty \leqq \mathsf{n}(\mathfrak{A}) \leqq \mathsf{m}(\mathfrak{A}) < +\infty, \qquad \mathsf{m}(\mathfrak{A}) \gtreqless 0$$

oder

$$(424)' \qquad -\infty < \mathsf{n}(\mathfrak{A}) \leqq \mathsf{m}(\mathfrak{A}) \leqq +\infty, \qquad \mathsf{n}(\mathfrak{A}) \gtreqless 0$$

gilt. In der Tat ist $\mathsf{n}(\mathfrak{A}_{[n]})$ der kleinste, $\mathsf{m}(\mathfrak{A}_{[n]})$ der größte Eigenwert von $\mathfrak{A}_{[n]}$, und nach Definition (S. 124) ist

$$\mathsf{n}(\mathfrak{A}) = \lim_{n=\infty} \mathsf{n}(\mathfrak{A}_{[n]}), \qquad \mathsf{m}(\mathfrak{A}) = \lim_{n=\infty} \mathsf{m}(\mathfrak{A}_{[n]})$$

zu setzen, wobei $+\infty$ und $-\infty$ als Grenzwerte zugelassen sind. Das Abschnittsspektrum braucht, wie bereits in (424), (424)' vermerkt ist, bei halbbeschränkten Matrizen ebensowenig auf derselben Seite des Nullpunktes zu liegen, wie bei den beschränkten (d. h. Abschnittseigenwerte von verschiedenem Vorzeichen werden nicht ausgeschlossen, $\mathfrak{A}$ braucht nicht einmal pseudodefinit zu sein).

Aus dem Toeplitzschen Satz gewinnt man den folgenden: Jede halbbeschränkte Matrix ist statistisch sinnvoll. — Denn es sei $\mathfrak{A}$ etwa von oben halbbeschränkt, so daß (424) zutrifft. Die Abschnittseigenwerte $\lambda_\nu^{(n)}$ von $\mathfrak{A}$ sind dann alle $\leqq \mathsf{m}(\mathfrak{A}) < +\infty$. Die Abschnittseigenwerte der Matrix $\lambda^* \mathfrak{E} - \mathfrak{A}$ sind nach S. 219 gleich $\lambda^* - \lambda_\nu^{(n)}$, also $\geqq \varepsilon > 0$, wenn die Zahl $\lambda^* \geqq \varepsilon + \mathsf{m}(\mathfrak{A})$ gewählt wird. Für hinreichend große λ^* ist

also $\lambda^* \mathfrak{E} - \mathfrak{A}$ positiv definit, also ist nach dem Toeplitzschen Satz $\lim_{n=\infty} R_{ik}^{(n)}(\lambda^*)$ gewiß vorhanden. In den auf S. 229 benutzten Bezeichnungen ist nämlich $\mathfrak{B} = \lambda^* \mathfrak{E} - \mathfrak{A}$, ${}_0\lambda_\nu^{(n)} = \lambda^* - \lambda_\nu^{(n)}$, $B_{ik}^{[n]} = R_{ik}^{(n)}(\lambda^*)$. Da jedoch der Grenzwert (420) auf einer Halbgeraden $\lambda^* > \text{const.}$ existiert, so wird er nach dem Vitalischen Satz für alle nicht reellen λ und außerdem für $\lambda > \mathsf{m}(\mathfrak{A})$ gewiß existieren. Denn für $\mu > \mathsf{m}(\mathfrak{A})$ ist jedes $\sigma_{ik}^{(n)}(\mu)$ konstant, $= \delta_{ik}$, da alle Eigenwerte $< \mathsf{m}(\mathfrak{A})$ sind, und andererseits ist $\int_{-\infty}^{+\infty} | d\sigma_{ik}^{(n)}(\mu) | \leqq 1$, so daß die Schlußbemerkung auf S. 105 angewandt werden kann. Damit ist gezeigt, daß die halbbeschränkten Matrizen gewiß statistisch sinnvoll sind.

Wir bezeichnen die Gesamtheit der Stellen μ_0, zu welchen es keine Umgebung $\mu_0 - \varepsilon < \mu < \mu_0 + \varepsilon$ gibt, auf welcher ein jedes $\overset{*}{\sigma}_{ik}(\mu)$ konstant wäre, als das Spektrum der statistisch sinnvollen Matrix. Aus $\sigma_{ik}^{(n)} [\rightarrow] \overset{*}{\sigma}_{ik}$ folgt wegen $\lambda_\nu^{(n)} \leqq \mathsf{m}(\mathfrak{A})$ offenbar, daß das Spektrum ebenfalls auf der Halbgeraden $\mu \leqq \mathsf{m}(\mathfrak{A})$ liegt. Es wäre aber denkbar, daß das Spektrum auf einer Halbgeraden liegt, ohne daß das Abschnittsspektrum es tut, d. h. ohne daß die Matrix halbbeschränkt ist. Dann wäre aber unser hinreichendes Kriterium des statistisch sinnvollen Charakters praktisch unbrauchbar. Z. B. bei wasserstoffähnlichen Spektren liegt das Spektrum der Energiematrix auf einer Halbgeraden, aber es ist eben nur das Spektrum und nicht das Abschnittsspektrum zugänglich. Ich beweise aber auch, daß das Abschnittsspektrum einer Q-Matrix auf einer Halbgeraden liegen muß, wenn das Spektrum auf einer Halbgeraden liegt. Die wasserstoffähnlichen Spektren gehören daher gewiß zu halbbeschränkten, also erst recht zu statistisch sinnvollen Matrizen.

§ 109. Ein Satz über das Abschnittsspektrum der halbbeschränkten Matrizen.

Wir beweisen zunächst, daß der *Endpunkt* des Abschnittsspektrums einer halbbeschränkten, der Q-Bedingung nicht notwendig genügenden Matrix auch dem Spektrum angehört, so daß z. B. im Fall (424) der Punkt $\mu = \mathsf{m}(\mathfrak{A})$ des Abschnittsspektrums zugleich ein Punkt des Spektrums ist. Wie wenig selbstverständlich dies ist, geht aus der bei (246) erwähnten Erscheinung hervor.

Beim Beweis gehen wir davon aus, daß für eine positiv definite n-äre Matrix $\mathfrak{C}_{(n)}$ und ihre Reziproke $\mathfrak{C}_{(n)}^{-1}$ gilt: $1 = \mathsf{n}(\mathfrak{C}_{(n)})\, \mathsf{m}(\mathfrak{C}_{(n)}^{-1})$, da

die Eigenwerte (>0) von $\mathfrak{C}_{(n)}$ die reziproken Eigenwerte (>0) von $\mathfrak{C}_{(n)}^{-1}$ sind, so daß der kleinste Eigenwert $\mathsf{n}(\mathfrak{C}_{(n)})$ von $\mathfrak{C}_{(n)}$ der reziproke Wert des größten Eigenwertes $\mathsf{m}(\mathfrak{C}_{(n)}^{-1})$ von $\mathfrak{C}_{(n)}^{-1}$ sein muß. Wir nehmen an, daß der Fall (424) vorliegt und setzen

$$(425) \qquad \lambda^* = \mathsf{m}(\mathfrak{A}) + \varepsilon, \qquad \varepsilon > 0.$$

Die Eigenwerte von $\mathfrak{A}_{[n]}$ sind $\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}$, also sind die Eigenwerte von $(\lambda^* \mathfrak{E} - \mathfrak{A})_{[n]}$ offenbar $\lambda^* - \lambda_1^{(n)}, \ldots, \lambda^* - \lambda_n^{(n)}$. Die größte unter den n Zahlen $\lambda_\nu^{(n)}$ ist $= \mathsf{m}(\mathfrak{A}_{[n]}) \leqq \mathsf{m}(\mathfrak{A}) < \lambda^*$, so daß die Eigenwerte von $\lambda^* \mathfrak{E}_{[n]} - \mathfrak{A}_{[n]}$ durchweg positiv sind, und zwar ist der kleinste Eigenwert, nämlich $\mathsf{n}(\lambda^* \mathfrak{E}_{[n]} - \mathfrak{A}_{[n]})$, gleich $\lambda^* - \mathsf{m}(\mathfrak{A}_{[n]})$. Wenn nun $\mathfrak{C}_{(n)} = \lambda^* \mathfrak{E}_{[n]} - \mathfrak{A}_{[n]}$ gesetzt wird, so ist $\mathfrak{C}_{(n)}^{-1} = \| R_{ik}^{(n)}(\lambda^*) \|_{(n)}$, es folgt daher

$$1 = [\lambda^* - \mathsf{m}(\mathfrak{A}_{[n]})] \; \mathsf{m}(\| R_{ik}^{(n)}(\lambda^*) \|_{(n)}),$$

da ja $1 = \mathsf{n}(\mathfrak{C}_{(n)})\, \mathsf{m}(\mathfrak{C}_{[n]}^{-1})$ gilt. Nimmt man hierbei den Grenzübergang $n \to \infty$ vor, so bleibt, da nach Definition $\mathsf{m}(\mathfrak{A}_{[n]}) \to \mathsf{m}(\mathfrak{A}) < +\infty$ gilt, wegen (425) offenbar

$$(426) \qquad 1 = \varepsilon \lim_{n=\infty} \mathsf{m}(\| R_{ik}^{(n)}(\lambda^*) \|_{(n)}).$$

Nun ist $\| R_{ik}^{(n)}(\lambda^*) \|_{(n)}$ die Reziproke des n-ten Abschnittes der positiv definiten Matrix $\lambda^* \mathfrak{E} - \mathfrak{A}$, so daß nach dem noch zu beweisenden Toeplitzschen Satz (siehe S. 229) $R_{ik}^{(n)}(\lambda^*) \to R_{ik}(\lambda^*)$ gilt, wobei $\| R_{ik}(\lambda^*) \|$ eine beschränkte Matrix bezeichnet. Beim Beweise des Toeplitzschen Satzes wird sich herausstellen, daß

$$(427) \qquad \lim_{n=\infty} \mathsf{m}(\| R_{ik}^{(n)}(\lambda^*) \|_{(n)}) = \mathsf{m}(\| R_{ik}(\lambda^*) \|)$$

gilt. Setzt man dies in (426) ein und nimmt sodann den Grenzübergang $\varepsilon \to +0$ vor, so folgt

$$(428) \qquad \lim_{\varepsilon \to +0} \mathsf{m}(\| R_{ik}(\lambda^*) \|) = +\infty.$$

Wegen (425) können wir dies so ausdrücken: Die Grenzresolvente $\| \lim R_{ik}^{(n)}(\lambda) \| = \| R_{ik}(\lambda) \|$ ist wohl für $\lambda > \mathsf{m}(\mathfrak{A})$ beschränkt, sie wird aber nicht beschränkt, wenn $\lambda \to \mathsf{m}(\mathfrak{A}) + 0$ geht. Daraus folgt nun leicht der zu beweisende Satz, daß es nämlich keine Umgebung

$$\mathsf{m}(\mathfrak{A}) - \eta < \mu < \mathsf{m}(\mathfrak{A}) + \eta$$

des Punktes $\mu = \mathsf{m}(\mathfrak{A})$ geben kann, auf welcher jedes Element der Grenzspektralmatrix $\| \overset{*}{\sigma}_{ik}(\mu) \|$ konstant wäre, daß also $\mu = \mathsf{m}(\mathfrak{A})$ dem Spektrum angehört.

Zwecks Zurückführung ad absurdum nehmen wir die Existenz eines solchen $\eta > 0$ an und bemerken, daß die Darstellung (421) gewiß auch für $\lambda > \mathsf{m}(\mathfrak{A})$ gilt. Vgl. nämlich die Schlußbemerkung auf S. 105. Nun ist $\sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \overset{*}{\sigma}_{ik}(\mu)\, x_i \bar{x}_n$ auf E nicht abnehmend, $\geqq 0$, $\leqq 1$. Somit ist, wenn λ nicht im Spektrum liegt, nach dem Hellyschen Satz über gliedweise Integration die Kopplungsform

$$\sum_{k=1}^{\infty} \sum_{i=1}^{\infty} R_{ik}(\lambda)\, x_i \bar{x}_k \quad \text{gleich} \quad \int_{-\infty}^{+\infty} \frac{d \sum\limits_{i=1}^{\infty} \sum\limits_{k=1}^{\infty} \overset{*}{\sigma}_{ik}(\mu)\, x_i \bar{x}_k}{\lambda - \mu},$$

also, wenn es ein $\eta \geqq 0$ geben würde, absolut

$$\leqq \int_{-\infty}^{+\infty} \frac{d \sum\limits_{i=1}^{\infty} \sum\limits_{k=1}^{\infty} \overset{*}{\sigma}_{ik}(\mu)\, x_i \bar{x}_k}{\eta} \leqq \eta^{-1},$$

sobald nur λ von $\mathsf{m}(\mathfrak{A})$ um weniger als η entfernt ist. Es bliebe also $\| R_{ik}(\lambda) \|$ für $\lambda \to \mathsf{m}(\mathfrak{A}) + 0$ beschränkt, im Widerspruch mit (428).

§ 110. Die wasserstoffähnlichen Spektra.

Es ist bereits erwähnt worden, daß das Spektrum der wasserstoffähnlichen Energiematrizen auf einer Halbgeraden liegt. Das Spektrum können wir dabei als die Gesamtheit derjenigen Stellen μ_0 definieren, zu welchen es kein $\varepsilon > 0$ gibt derart, daß in dem Intervalle $|\mu - \mu_0| < \varepsilon$ jedes $\sigma_{ik}(\mu)$ konstant, gleich $\sigma_{ik}(\mu_0)$ ist, wobei $\| \sigma_{ik}(\mu) \|$ *irgendeine* der Grenzspektralmatrizen bedeutet. Diese Definition ist (vorübergehend) notwendig, da wir zunächst noch nicht wissen, ob es nicht mehrere Grenzspektralmatrizen gibt. Nun sind die erwähnten Energiematrizen durchweg Q-Matrizen. Wir wollen daraus schließen, daß auch das Abschnittsspektrum auf einer Halbgeraden liegt, d. h. daß diese Matrizen halbbeschränkt, also nach § 108 gewiß statistisch sinnvoll sind.

Beim Beweis können wir ohne Einschränkung der Allgemeinheit annehmen, daß das soeben definierte Spektrum der Q-Matrix $\| a_{ik} \|$ auf der Halbgeraden $\mu > 0$ liegt, so daß nach der Schlußbemerkung von § 107

$$a_{ik} = \int_{-\infty}^{+\infty} \mu \, d\sigma_{ik}(\mu) = \int_{0}^{+\infty} \mu \, d\sigma_{ik}(\mu)$$

gilt. Die Behauptung, daß nämlich die Matrix $\| a_{ik} \|$ halbbeschränkt ist,

läuft, wie wir wissen, darauf hinaus, daß die Hermitesche Form

$$(*) \qquad \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik}\, x_i\, \bar{x}_k = \sum_{i=1}^{m}\sum_{k=1}^{m}\int_0^{+\infty} \mu\, d\,\sigma_{ik}(\mu)\, x_i\, \bar{x}_k$$

nichtnegativ definit ist. Nun ist aber

$$\sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \sigma_{ik}(\mu)\, x_i\, \bar{x}_k$$

eine nicht abnehmende Funktion von μ (vgl. S. 226). Setzt man hier $x_i = 0$ für $i > m$, so folgt, daß auch die Abschnittsform

$$\sum_{i=1}^{m}\sum_{k=1}^{m} \sigma_{ik}(\mu)\, x_i\, \bar{x}_k$$

nicht abnehmend, und daher der Ausdruck

$$\int_0^{+\infty} \mu\, d \sum_{i=1}^{m}\sum_{k=1}^{m} \sigma_{ik}(\mu)\, x_i\, \bar{x}_k$$

gewiß $\geqq 0$, d. h. die Form (*) nichtnegativ definit ist.

Damit ist gezeigt, daß die Endpunkte des Spektrums und des Abschnittsspektrums identisch sind, daß also jedes Hydrogenspektrum zu einer halbbeschränkten, also erst recht statistisch sinnvollen Matrix gehört (sobald die Q-Bedingung als erfüllt angenommen wird). — Unnötig ist zu betonen, daß dabei nicht verlangt wird, daß die volle Halbgerade $\mu \leqq \mathsf{m}(\mathfrak{A})$ bzw. $\mu \geqq \mathsf{n}(\mathfrak{A})$ vom Abschnittsspektrum oder vom Spektrum bedeckt sein soll, so daß Spektrumslücken (z. B. diejenigen zwischen den Balmertermen) sehr wohl mit erlaubt sind.

§ 111. Der Toeplitzsche Satz über definite Matrizen.

Wir haben jetzt noch den Beweis des seit S. 229 benutzten Toeplitzschen Satzes, sowie der Grenzgleichung (427) nachzuholen. Es sei also $\mathfrak{B} = \| b_{ik} \|$ eine Hermitesche Matrix, für welche die Bedingung (423) erfüllt ist, und $\mathfrak{B}_{[n]}^{-1} = \| B_{ik}^{[n]} \|_{(n)}$ die gewiß vorhandene Reziproke von $\mathfrak{B}_{[n]}$. Da $\mathfrak{B}_{[n]}$ positiv definit ist, so gibt es nach § 28 eine rekursive Matrix $\mathfrak{S}_{(n)}$ mit $\mathfrak{B}_{[n]} = \mathfrak{S}_{(n)}\mathfrak{S}_{(n)}^*$. Und zwar ist die unter (96) angegebene Matrix $\mathfrak{S}_{(n)}$ aus $\mathfrak{B}_{[n]}$ offenbar derart abgeleitet, daß der n-te Abschnitt von $\mathfrak{S}_{(n+1)}$ gleich $\mathfrak{S}_{(n)}$ ist. M. a. W., es gibt eine von n unabhängige unendliche rekursive Matrix $\mathfrak{S}$, deren n-ter Abschnitt $\mathfrak{S}_{[n]} = \mathfrak{S}_{(n)}$ ist. Man hat daher $\mathfrak{B}_{[n]} = \mathfrak{S}_{(n)}\mathfrak{S}_{(n)}^*$, also

$$\mathfrak{B}_{[n]}^{-1} = \mathfrak{S}_{(n)}^{*-1}\,\mathfrak{S}_{(n)}^{-1} = \mathfrak{S}_{[n]}^{-1*}\,\mathfrak{S}_{[n]}^{-1}.$$

Nun ist $\mathfrak{S}_{[n]}$ rekursiv, so daß in der Diagonale von Null verschiedene Elemente stehen, während oberhalb der Diagonale alles verschwindet. Die Reziproke $\mathfrak{S}_{[n]}^{-1}$ wird wieder rekursiv sein. Die zugehörigen homogenen Gleichungen können dabei eben *rekursiv* gelöst werden. Es versteht sich also von selbst, daß es wieder eine von n unabhängige unendliche rekursive Matrix $\mathfrak{T} = \| t_{ik} \|$ gibt derart, daß $\mathfrak{S}_{[n]}^{-1} = \mathfrak{T}_{[n]}$, mithin $\mathfrak{B}_{[n]}^{-1} = \mathfrak{T}_{[n]}^{*} \mathfrak{T}_{[n]}$ gilt, welche Matrix von $(\mathfrak{T}^{*} \mathfrak{T})_{[n]}$ wohl zu unterscheiden ist. Es zeigt sich nun sofort, daß in einem leicht verständlichen Sinne $\lim\limits_{n=\infty} \mathfrak{T}_{[n]}^{*} \mathfrak{T}_{[n]} = \lim\limits_{n=\infty} (\mathfrak{T}^{*} \mathfrak{T})_{[n]}$ gilt. Zunächst hatten wir nämlich $\mathfrak{B}_{[n]}^{-1} = \| B_{ik}^{[n]} \|_{(n)}$ gesetzt, es ist also

$$(429) \qquad B_{ik}^{[n]} = \sum_{j=1}^{n} \bar{t}_{ji} t_{jk}.$$

Man hat ferner wegen (53), (211), (213)

$$(\mathsf{P}(\mathfrak{T}_{[n]}))^2 = \mathsf{M}(\mathfrak{T}_{[n]}^{*} \mathfrak{T}_{[n]}) = \mathsf{M}(\mathfrak{B}_{[n]}^{-1}) = \mathsf{m}(\mathfrak{B}_{[n]}^{-1}) = \frac{1}{\mathsf{n}(\mathfrak{B}_{[n]})} \leqq \frac{1}{\mathsf{n}(\mathfrak{B})} = \text{konst.},$$

so daß $\mathfrak{T}$ beschränkt, also eine Q-Matrix ist, und daher die Reihe $\sum\limits_{j=1}^{\infty} \bar{t}_{ji} t_{jk}$, das (i, k)-te Element der beschränkten Matrix $\| B_{ik} \| = \mathfrak{T}^{*} \mathfrak{T}$, konvergiert. Wegen (429) hat man also $B_{ik}^{[n]} \to B_{ik}$, womit der Toeplitzsche Satz bewiesen ist. — Um endlich auch (427) zu beweisen, haben wir zu zeigen, daß

$$\lim_{n=\infty} \mathsf{m}(\| B_{ik}^{[n]} \|_{(n)}) = \lim_{n=\infty} \mathsf{m}(\| B_{ik} \|_{[n]})$$

gilt. Da es sich um positiv definite Matrizen handelt, können wir beidemal M an Stelle von m setzen. Es bleibt demnach nur zu beweisen übrig, daß

$$\lim_{n=\infty} \mathsf{M}(\| B_{ik}^{[n]} \|_{(n)}) = \mathsf{M}(\| B_{ik} \|),$$

d. h.

$$\lim_{n=\infty} \mathsf{M}(\mathfrak{T}_{[n]}^{*} \mathfrak{T}_{[n]}) = \mathsf{M}(\mathfrak{T}^{*} \mathfrak{T}),$$

oder mit Rücksicht auf (53), (235) einfach $(\mathsf{P}(\mathfrak{T}_{[n]}))^2 \to (\mathsf{P}(\mathfrak{T}))^2$ gilt, was aber per definitionem (S. 124) richtig ist.

Es ist bereits vermerkt worden, daß die Halbbeschränktheit zwar hinreichend, aber nicht notwendig ist für den statistisch sinnvollen Charakter. Die Behauptung kann am einfachsten mit der Toeplitzschen Theorie der Potenzreihen $f(z) = \sum\limits_{n=0}^{\infty} \alpha_{n+1} z^n$ vom Konvergenzradius 1

belegt werden. Es seien die Koeffizienten α reell, also die Matrix $\| \alpha_{|i-k|} \|$ Hermitesch. Es sei ferner $\sum_{n=1}^{\infty} |\alpha_n|^2$ konvergent, d. h. $f(z)$ möge beim radialen Grenzübergang $|z| \to 1$ in eine (bis auf eine Nullmenge eindeutig bestimmte), samt ihrem Quadrat nach Lebesgue integrierbare Randfunktion $f(e^{i\theta})$ übergehen, ohne daß $\| \alpha_{|i-k|} \|$ auch beschränkt ist. Der Cauchyschen Integralformel (in welche hier die Integraldarstellung der Resolvente übergeht), läßt sich dann nach den Elementen der Funktionentheorie entnehmen, daß $\| \alpha_{|i-k|} \|$ statistisch sinnvoll ist, indem die Grenzwerte (420) existieren. Andererseits können wir die Randfunktion $f(e^{i\theta})$ derart wählen, daß sie, abgesehen von etwa endlich vielen logarithmischen Unendlichkeitsstellen, für $0 \leqq \theta \leqq 2\pi$ analytisch ist und ihr Wertevorrat die ganze reelle Achse der f-Ebene bedeckt. Dann ist $\| \alpha_{|i-k|} \|$ gewiß nicht halbbeschränkt und dennoch statistisch sinnvoll. Doch ist nicht jede Hermitesche Q-Matrix statistisch sinnvoll: der auf S. 221 unter c) erwähnte Fall ist, wie wir jetzt sehen werden, bei passendem $\mathfrak{A}$ sehr wohl möglich.

§ 112. Das Seidel-M. Sternsche Kriterium bei speziellen Jacobischen Matrizen. Über die Koordinatenmatrix des linearen Oszillators.

Eine der Bedingung

$$(430) \qquad a_{ik} = 0, \quad |i-k| > 1; \qquad a_{i\,i+1} = \bar{a}_{i+1\,i} \neq 0; \qquad a_{ii} \gtreqless 0$$

genügende (ersichtlich Hermitesche) Matrix, die sonst beliebig sein kann, bezeichnen wir allgemein als eine Jacobische Matrix [sie ist, wie man durch Anschreiben der Kopplungsform von $\mathfrak{A}_{[n]}$ sofort erkennt, dann und nur dann beschränkt, wenn $|a_{ii}| <$ const., $|a_{i\,i+1}| <$ Const. gilt, d. h. wenn die Elemente beschränkt sind; bei beliebigen Matrizen ist es freilich nicht so]. Die Jacobischen Matrizen sind finit, also gewiß Q-Matrizen.

Ist in einer Jacobischen Matrix $\mathfrak{A}$ auch jedes Diagonalelement a_{ii} gleich Null, so daß (vgl. S. 70)

$$(431) \qquad \mathfrak{A} = \left\| \begin{matrix} 0 & -b_1 & 0 & 0 & \dots \\ -\bar{b}_1 & 0 & -b_2 & 0 & \dots \\ 0 & -\bar{b}_2 & 0 & -b_3 & \dots \\ 0 & 0 & -\bar{b}_3 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{matrix} \right\|; \qquad a_{i\,i+1} = -b_i \neq 0$$

wird, so sind die $R_{11}^{(n)}(\lambda)$, d. h. die „Eckenelemente“ in der Resolvente

der $\mathfrak{A}_{[n]}$, nach S. 73 die Näherungsbrüche des formal gebildeten unendlichen Kettenbruches

(432)
$$\varkappa(\lambda) = \cfrac{-1}{-\lambda - \cfrac{|b_1|^2}{-\lambda - \cfrac{|b_2|^2}{-\lambda - \cdots}}}.$$

Dieser Kettenbruch ist aber nach Seidel und M. Stern für rein imaginäre λ dann und nur dann konvergent, wenn die Reihe

(433)
$$\sum_{n=1}^{\infty} B_n, \quad \text{wo} \quad B_n = \frac{1}{|b_n|^2 B_{n-1}}, \quad B_0 = 1$$

ist, divergiert. Ist also die Reihe (433) konvergent, was z. B. für $\| b_n \| = C n^\delta$, $\delta > 1$ zutrifft, so ist der Grenzwert $\lim\limits_{n=\infty} R_{ik}^{(n)}(\lambda)$ für $i = k = 1$ und für rein imaginäre λ nicht vorhanden, also ist dann (vgl. S. 225) die Matrix (431) gewiß nicht statistisch sinnvoll. Es ließe sich leicht zeigen, daß die Divergenz der Reihe (433) nicht nur notwendig, sondern auch hinreichend ist dafür, daß (431) sich als statistisch sinnvoll erweist. Da die Reihe für $|b_n| = C n^\delta$, $\delta \leqq 1$ divergiert, so ist daher die wohlbekannte Koordinatenmatrix des Heisenbergschen linearen Oszillators gewiß statistisch sinnvoll [jedoch, wie aus der vorher gemachten Bemerkung hervorgeht, gewiß nicht beschränkt]. Unter Benutzung der auf S. 194 mit $\chi_n(\lambda)$ bezeichneten Polynome, die übrigens die Näherungsnenner von (432) sind, kann man die Heisenbergsche Matrix sofort in die Schrödingersche Fassung umrechnen. Bei wirklichen physikalischen Problemen kommen freilich Matrizen mit mehreren Zeigerpaaren in Betracht und durch die Separation der Variablen kann, wie das Wasserstoffmodell zeigt, der formale „Jacobische" Charakter der Matrix verloren gehen.

§ 113. Beliebige Jacobische Matrizen.

Das Divergenzkriterium (433) entscheidet nur, ob der Fall c) oder aber nicht der Fall c) vorliegt (S. 221), kann aber, wenn der Fall c) nicht vorliegt, nicht entscheiden, welche der beiden weiteren Möglichkeiten a), b) vorliegt. Nun gibt es aber ein Kriterium, welches nicht nur im Falle (431), sondern bei beliebigen Jacobischen Matrizen (430) zu entscheiden gestattet, ob der Fall a) oder aber nicht der Fall a) vorliegt.

Eine beliebige Jacobische Matrix ist, wie sich ebenso wie in dem Spezialfalle (432) zeigen läßt, dann und nur dann statistisch sinnvoll, wenn der „assoziierte Kettenbruch“

$$\varkappa(\lambda) = \cfrac{-1}{a_1 - \lambda - \cfrac{|b_1|^2}{a_2 - \lambda - \cfrac{|b_2|^2}{a_3 - \lambda - \cfrac{|b_3|^2}{a_3 - \lambda - \cdots}}}} \tag{434}$$

für alle nicht reellen λ, oder, was nach dem Satze von Vitali dasselbe ist, etwa für alle rein imaginären λ konvergiert; und zwar hat man dann in jedem nicht reellen Bereiche sogar eine gleichmäßige Konvergenz. Dies reicht jedoch noch nicht aus, um die Regularität der Jacobischen Matrix, d. h. das Vorliegen des Falles a) sicherzustellen. Hierfür muß vielmehr $\varkappa(\lambda)$ einen von Hamburger entdeckten besonderen Konvergenzcharakter aufweisen, der von Hellinger mit den Weylschen Untersuchungen über lineare Randwertprobleme bei wesentlich singulären Differentialgleichungen in Zusammenhang gebracht wurde. Dieser besondere Konvergenzcharakter, die „vollständige Konvergenz“, folgt aus der einfachen gleichmäßigen Konvergenz noch nicht. Es gibt also statistisch sinnvolle Matrizen, die nicht regulär sind: der Fall b) auf S. 221 ist keine leere Möglichkeit. Wir wollen über das Regularitätsproblem der allgemeinen Jacobischen Matrizen ebenfalls nähere Angaben machen, doch müssen wir uns leider auf ein Referat beschränken und wegen Beweis auf die Arbeiten von Hamburger und Hellinger, sowie auch auf die mehrfach erwähnte Schrift von Carleman verweisen, der diese Ergebnisse in seine allgemeine Theorie eingeordnet hat.

§ 114. Das Hamburgersche Momentenproblem.

Als Momentenproblem bezeichnet man die Aufgabe, das unendliche System der Integralgleichung

$$\int_{-\infty}^{+\infty} \mu^n \, d\varrho(\mu) = \mathrm{m}_n; \qquad n = 0, 1, 2, \ldots, \tag{435}$$

bei gegebenen m_n durch eine *nicht abnehmende* Funktion $\varrho(\mu)$ zu erfüllen, die *keine Treppenfunktion* sein soll, also entweder unendlich viele Sprünge oder ein Nichtkonstanzintervall haben muß. Offenbar müssen

alle $\mathfrak{m}_n$ reell sein. Hat das Momentenproblem eine Lösung, so sind alle Integrale (435) absolut konvergent. Denn die $\mathfrak{m}_{2n}$ sind sicherlich absolut konvergent, und für die ungeraden Momente gilt nach der Schwarzschen Ungleichheit

$$(436) \qquad \int_{-\infty}^{+\infty} |\mu|^{2n+1}\, d\varrho(\mu) \leqq \sqrt{\mathfrak{m}_{2n}\,\mathfrak{m}_{2n+2}}\,.$$

Da es sich um Stieltjessche Integrationen handelt, haben wir zwei Lösungen, die — bis auf eine additive Konstante — an ihren Stetigkeitsstellen übereinstimmen, nicht als verschieden zu betrachten, so daß wir fortan nur Lösungen zulassen wollen, die von rechts stetig sind und in denen die additive Konstante durch $\varrho(-\infty) = 0$ festgelegt ist. Letztere Normierung ist stets möglich. Denn gibt es überhaupt eine Lösung, so ist sie wegen $\int_{-\infty}^{+\infty} d\varrho(\mu) = \mathfrak{m}_0$ gewiß beschränkt, sie ist ferner nicht abnehmend. Man hat daher $\varrho(+\infty) - \varrho(-\infty) = \mathfrak{m}_0$, so daß die Normierung auf $\varrho(+\infty) = \mathfrak{m}_0$ hinausläuft. Es wurde bereits betont, daß es beschränkte, nicht abnehmende Funktionen gibt, für welche die Integrale (435) von $n = 1$ an divergieren. — Umgekehrt ist auch (435) nicht immer lösbar: es wird von der Lösung zu viel verlangt. Damit das Momentenproblem lösbar ist, müssen, wie sich leicht zeigen läßt, die Abschnittseigenwerte der unendlichen Hankelschen Matrix[1])

$$(437) \qquad \|\mathfrak{m}_{i+k}\| = \left\| \begin{matrix} \mathfrak{m}_0 & \mathfrak{m}_1 & \mathfrak{m}_2 & \cdot \\ \mathfrak{m}_1 & \mathfrak{m}_2 & \cdot & \cdot \\ \mathfrak{m}_2 & \cdot & \cdot & \\ \cdot & \cdot & & \end{matrix} \right\|$$

durchweg > 0 sein (sie dürfen sich bei Null häufen, nicht aber auch verschwinden, da sonst die Lösung höchstens eine Treppenfunktion sein könnte). Hamburger hat nun bewiesen, daß diese Bedingung für die Lösbarkeit des Momentenproblems (435) nicht nur notwendig,

[1]) Die (reellen) Toeplitzschen Matrizen (vgl. S. 208) sehen demgegenüber so aus:

$$\|\alpha_{i-k}\| = \left\| \begin{matrix} \alpha_0 & \alpha_1 & \alpha_2 & \cdot \\ \alpha_1 & \alpha_0 & \alpha_1 & \cdot \\ \alpha_2 & \alpha_1 & \alpha_0 & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{matrix} \right\|.$$

sondern auch hinreichend ist. Hamburger führt den Beweis seines Existenzsatzes in der Sprache der Kettenbrüche. Einen mehr direkten Beweis hat später im Anschluß an einen von Stridsberg gegebenen Algorithmus M. Riesz gegeben. Der springende Punkt ist beidemal der Hellysche Fortpflanzungs- bzw. Auswahlsatz von Kap. II.

§ 115. Fortsetzung. Bestimmte Momentenprobleme.

Die Frage, wann das Momentenproblem mindestens eine Lösung hat, ist damit erledigt. Uns interessiert indessen eine andere Frage: wann ist das Momentenproblem *bestimmt* (déterminé)? M. a. W., gesetzt, daß die Hamburgersche Bedingung erfüllt, d. h. mindestens eine Lösung ϱ vorhanden ist, wann gibt es *nur* eine Lösung? [Zwei Lösungen ϱ_1, ϱ_2, für welche $\varrho_1(\mu+0)-\varrho_2(\mu+0)\equiv$ konst. gilt, sind dabei nicht als verschieden zu betrachten.] Die Antwort hat Hamburger ebenfalls in der Sprache der Kettenbruchlehre geliefert (siehe oben, S. 238). Den Zusammenhang seines Ergebnisses mit der Theorie der unendlichen Matrizen hat, wie erwähnt, Hellinger aufgedeckt. Wir gehen nunmehr zu diesen Untersuchungen über.

Es sei das Momentenproblem lösbar und es sei $\varrho(\mu)$ eine Lösung. Man kann dann eine (und, bis auf den belanglosen Faktor ± 1, auch nur eine) Folge von reellen Polynomen $\chi_1, \chi_2, \ldots$, wobei χ_n genau vom Grade $n-1$ ist, derart bestimmen, daß

$$\int_{-\infty}^{+\infty} \chi_i(\mu)\,\chi_k(\mu)\,d\varrho(\mu)=\delta_{ik} \tag{438}$$

wird (vgl. S. 188). Dieses Integral, und allgemeiner das Integral

$$\int_{-\infty}^{+\infty} \Phi(\mu)\,d\varrho(\mu) \qquad (\Phi = \text{ein Polynom}) \tag{439}$$

ist gewiß absolut konvergent, da ja die Momente (435) existieren. Ferner hat das Integral (439) für alle evtl. verschiedenen Lösungen ϱ von (435) einen und denselben Wert, da ja eben alle diese ϱ Lösungen von (435) sind. Man kann also die vollkommen bestimmte Matrix

$$\| a_{ik} \| = \left\| \int_{-\infty}^{+\infty} \mu\,\chi_i(\mu)\,\chi_k(\mu)\,d\varrho(\mu) \right\| \tag{440}$$

bilden, welche offenbar (vgl. S. 189) eine Jacobische Matrix ist. Es stellt sich nun heraus, daß das (lösbare) Momentenproblem (435) dann und nur dann bestimmt ist, wenn diese Jacobische Matrix regulär ist, d. h.

wenn für (440) der Fall a) von S. 221 vorliegt. — Umgekehrt kann jede Jacobische Matrix (430) in der Gestalt (440) dargestellt werden. Die Polynome χ_i sind dabei immer restlos bestimmt, die Belegung hingegen dann und nur dann, wenn die mit ihr gebildeten Zahlen $\mathfrak{m}_n$ zu einem *bestimmten* Momentenproblem gehören.

Ein hinreichendes Kriterium dafür, daß das als lösbar vorausgesetzte Momentenproblem (435) bestimmt ist, besteht nach Carleman in der Divergenz der Reihe

(441) $$\sum_{n=0}^{\infty} \frac{1}{\sqrt[2n]{\mathfrak{m}_{2n}}}.$$

Es genügt also, wenn $|\sqrt[n]{\mathfrak{m}_n}| < K n^2$ ist. Dies ist bereits von Hamburger und von ihm unabhängig von Pólya gefunden worden. Hamburger hat zugleich gezeigt, daß seine Bedingung nicht wesentlich weiter verschärft werden kann, indem es lösbare und nicht bestimmte Momentenprobleme gibt, für welche $|\sqrt[n]{\mathfrak{m}_n}| < K n^{2+\varepsilon}$ bei jedem $\varepsilon > 0$ gilt. Vgl. auch die Bemerkung auf S. 102.

Setzt man

(442) $$\varrho(\mu) \equiv 0, \quad \mu < 0; \quad \varrho(\mu) = \int_0^{\mu} (1 - \sin \sqrt[4]{\nu})\, e^{-\sqrt[4]{\nu}}\, d\nu, \quad \mu \geqq 0,$$

so ist das Momentenproblem nicht bestimmt[1]), d. h. es gibt von (442) wesentlich verschiedene Belegungen, die bez. dieselben Momente wie (442) haben, so daß (440) nicht regulär, wohl aber halbbeschränkt, also statistisch sinnvoll ist. Der Beweis ergibt sich aus der Theorie von Stieltjes, der in einer unsterblichen Abhandlung alle diese Fragen überhaupt aufgeworfen, seinen Integralbegriff eingeführt, das Momentenproblem in dem nach ihm benannten Spezialfall $\varrho(\mu) \equiv 0$, $\mu < 0$ samt den zugehörigen Jacobischen Matrizen eigentlich restlos erledigt hat — obwohl dies erst später, eigentlich nur zwanzig Jahre nach Stieltjes zutage trat, als nämlich Toeplitz und Hellinger die Zusammenhänge mit der Hilbertschen Theorie aufgedeckt haben (vgl. § 87) — so daß Stieltjes wohl als der Vater der Spektraltheorie zu betrachten ist. Er legte dabei nicht die „assoziierten Kettenbrüche" (434), sondern die „korrespondierenden" zugrunde, die eben dem Stieltjesschen Spezialfall angepaßt und dementsprechend u. U. bereits dann versagen, wenn

[1]) Denn es gilt

$$\int_0^{+\infty} \mu^n e^{-\sqrt[4]{\mu}} \sin \sqrt[4]{\mu}\, d\mu = 0 \quad \text{für} \quad n = 0, 1, 2, \ldots.$$

$\varrho(\mu) \equiv 0$, $\mu < c$ nur mit $c < 0$ gilt, trotzdem dabei die Matrix halbbeschränkt ist. — Ist das Stieltjessche Momentenproblem

$$\int_0^{+\infty} \mu^n \, d\varrho(\mu) = \mathrm{m}_n; \qquad n = 0, 1, 2, \ldots \tag{443}$$

nicht *bestimmt*, so ist es erst recht das Hamburgersche nicht, das zu denselben m_n gehört. Man setze nämlich die zunächst nur für $\mu \geqq 0$ erklärte Lösung von (443) für $\mu < 0$ einfach gleich dem Wert, welche die Lösung von (443) für $\mu = 0$ annimmt.

Da unser $\Phi(\|\sigma_{ik}(\mu)\|)$ nicht abnehmend ist, so läßt (305) unmittelbar erraten, daß das Momentenproblem nicht nur für die Jacobischen Matrizen von Bedeutung ist. Allerdings haben wir jetzt zu beachten, daß, wenn $\mathfrak{A}$ nicht beschränkt ist, die Potenzen von $\mathfrak{A}$ nicht zu existieren brauchen. Ist aber $\mathfrak{A}$ finit, so ist auch $\mathfrak{A}^n$ finit, also eine Q-Matrix. Es sei allgemeiner $\mathfrak{A}$ eine Hermitesche Q-Matrix, deren sämtliche Potenzen vorhanden und außerdem Q-Matrizen sind. Dann kann man nach Carleman nach einer sinngemäßen Wiederholung der § 107 angewandten Schlußweisen zeigen, daß (305) gewiß gilt, wobei $\|\sigma_{ik}\|$ die Grenzspektralmatrix der etwa als statistisch sinnvoll angenommenen Matrix $\mathfrak{A}$ bezeichnet.

§ 116. Die Carlemansche „Feldtheorie". Statistische Stabilität.

Wir haben gesehen, daß es Momentenprobleme gibt, die nicht bestimmt sind, trotzdem die Jacobische Matrix statistisch sinnvoll ist. Es gibt also dann zwar nur eine *Grenz*spektralmatrix, es gibt aber noch weitere, von der Grenzspektralmatrix verschiedene, jedoch mit ihr an sich gleichberechtigte „Spektralmatrizen" für die Jacobische Matrix; nur treten sie durch Betrachtung der Abschnitte nicht in Evidenz. Es liegt also nahe, den Grenzprozeß, der den n-ten Abschnitt der unendlichen Matrix in die unendliche Matrix überführt, derart zu verallgemeinern, daß auch weitere Spektralmatrizen erscheinen.

Wir gehen nunmehr zu einer diesbezüglichen Untersuchung von Carleman über.

Wir denken uns alle Hermiteschen Q-Matrizen $\mathfrak{A}$ zu einem „Raum", oder wie wir lieber sagen wollen, zu einem „Feld" vereinigt, in welchem jedem „Punkte" eine und nur eine Matrix entspricht, und umgekehrt. Wir sagen, eine „Punktfolge", d. h. eine Folge $\mathfrak{A}^{\{1\}}, \mathfrak{A}^{\{2\}}, \ldots$ von unendlichen Hermiteschen Q-Matrizen $\mathfrak{A}^{\{n\}} = \|a_{ik}^{\{n\}}\|$ konvergiere im Mittel

gegen den Punkt $\mathfrak{A} = \|a_{ik}\|$, und schreiben hierfür kurz $\mathfrak{A}^{\{n\}}\{\rightarrow\}\mathfrak{A}$, $n \rightarrow +\infty$, wenn

$$(444) \qquad \lim_{n=\infty} \sum_{k=1}^{\infty} |a_{ik} - a_{ik}^{\{n\}}|^2 = 0$$

bei einem jeden i gilt. Es wird nicht verlangt, daß diese Grenzgleichung *gleichmäßig* in i besteht. Versteht man unter $\mathfrak{A}^{[n]} = \|a_{ik}^{[n]}\|$ die unendliche Matrix, deren n-ter Abschnitt mit dem n-ten Abschnitt von $\mathfrak{A}$ identisch ist, während $a_{ik}^{[n]} = \delta_{ik}$ gilt, wenn mindestens einer der beiden unteren Zeiger $> n$ ausfällt, so gilt $\mathfrak{A}^{[n]}\{\rightarrow\}\mathfrak{A}$ gewiß. Der Einführung der Grenzspektralmatrizen liegt ersichtlich die Beschränkung auf diese einfachsten, $\mathfrak{A}$ mittelkonvergent approximierenden Folgen zugrunde. Wir lassen jedoch jetzt nicht nur diese $\{\mathfrak{A}^{[n]}\}$, sondern beliebige Folgen $\{\mathfrak{A}^{\{n\}}\}$ mit $\mathfrak{A}^{\{n\}}\{\rightarrow\}\mathfrak{A}$ zu und fassen den Ersatz von $\mathfrak{A}$ durch $\mathfrak{A}^{\{n\}}$ und den darauf folgenden Grenzübergang $n \rightarrow \infty$ als eine Art virtuelle Störung von $\mathfrak{A}$ auf. Wir werden sehen, daß für $\mathfrak{A}$ durch diese Störungen ebenso Spektralmatrizen induziert werden, wie dies im Spezialfalle $\mathfrak{A}^{\{n\}} = \mathfrak{A}^{[n]}$ der Fall war (S. 225). Hatten dabei alle Teilfolgen von $\mathfrak{A}^{[n]}$ für $\mathfrak{A}$ dieselbe Grenzspektralmatrix induziert, so haben wir $\mathfrak{A}$ als statistisch sinnvoll bezeichnet. Hat die Matrix $\mathfrak{A}$ die wesentlich weitergehende Eigenschaft, daß *alle* mittelkonvergenten Störungen für $\mathfrak{A}$ eine und dieselbe verallgemeinerte, „feldmäßige“ Spektralmatrix induzieren, so werden wir dies dahin zu deuten haben, daß $\mathfrak{A}$ im Felde eine stabile Lage hat, kurz, daß $\mathfrak{A}$ nicht nur statistisch sinnvoll, sondern außerdem statistisch stabil ist [während der statistisch sinnvolle Charakter allein nur eine Art *bedingte* Stabilität, d. i. Stabilität *gewisser* Störungen gegenüber ausspricht]. — Ich versehe hierbei nur ein Ergebnis von Carleman mit einer anschaulichen Nomenklatur und muß jetzt zum mathematischen Inhalt übergehen, und vor allem erklären, was überhaupt bei beliebigen Störungen unter dem „Induzieren“ zu verstehen sei.

§ 117. Feldresolvente und Feldspektralmatrix.

Es sei $\mathfrak{A}^{\{n\}}\{\rightarrow\}\mathfrak{A}$ und es bezeichne $\overset{*}{\mathfrak{R}}{}_\lambda^{\{n\}}$ die Grenzresolvente von $\mathfrak{A}^{\{n\}}$, oder, wenn $\mathfrak{A}^{\{n\}}$ nicht statistisch sinnvoll ist, irgendeine der Grenzresolventen, endlich sei $\|\overset{*}{\sigma}{}_{ik}^{\{n\}}(\mu)\|$ eine zu $\mathfrak{A}^{\{n\}}$ gehörige Grenzspektralmatrix und zwar eben diejenige, für welche

$$(445) \qquad \overset{*}{\mathfrak{R}}{}_\lambda^{\{n\}} = \left\| \int_{-\infty}^{+\infty} \frac{d\overset{*}{\sigma}{}_{ik}^{\{n\}}(\mu)}{\lambda - \mu} \right\| \equiv \|\overset{*}{R}{}_{ik}^{\{n\}}(\lambda)\|$$

gilt. Unter λ ist durchweg irgendeine nicht reelle Zahl verstanden. Wegen (409) ist $\mathsf{M}(\|\overset{*}{R}{}_{ik}^{\{n\}}(\lambda)\|) \leqq \frac{1}{\mathfrak{J}[\lambda]}$. Ersetzt man hierbei M durch P, so multipliziert sich die Schranke wegen (212) höchstens mit 2, mithin ist (vgl. S. 130)

$$\sum_{i=1}^{\infty} |\overset{*}{R}{}_{ik}^{\{n\}}(\lambda)|^2 < \Gamma(\lambda) < +\infty, \tag{446}$$

wobei $\Gamma(\lambda)$ nur von λ abhängt. Wegen (444) sind die in (444) hinter dem Limeszeichen stehenden Summen unterhalb einer höchstens von i abhängigen, also von n unabhängigen Schranke gelegen. Da andererseits $\sum_{k=1}^{\infty} |a_{ik}|^2$ bei einem jeden i konvergiert ($\mathfrak{A}$ ist ja eine Q-Matrix), so folgt daraus, nach einer leichten Benutzung der Schwarzschen Ungleichung und ihrer Umkehrung, daß es eine von n unabhängige, also höchstens von i abhängige Schranke C_i geben muß, für welche

$$\sum_{k=1}^{\infty} |a_{ik}^{\{n\}}|^2 < C_i < +\infty \tag{447}$$

ausfällt, wobei freilich C_i von der mittelkonvergenten Folge $\{\mathfrak{A}^{\{n\}}\}$ abhängt, die wir fest gewählt haben.

Mit Rücksicht auf (421), wegen $\int_{-\infty}^{+\infty} |d\overset{*}{\sigma}{}_{ik}^{\{n\}}(\mu)| \leqq 1$ und nach dem Montelschen Satz können wir nun aus $\{\mathfrak{A}^{\{n\}}\}$ eine Teilfolge $\{\mathfrak{A}^{\{h_n\}}\}$ derart aussondern, daß $\overset{*}{R}{}_{ik}^{\{h_n\}}(\lambda)$ für $n \to \infty$ und für alle nicht reellen λ einer Grenze $\overset{**}{R}_{ik}(\lambda)$ zustrebt, wobei mit Rücksicht auf das Diagonalprinzip angenommen werden kann, daß die h_n von (i, k) unabhängig sind. Die Matrix $\overset{**}{\mathfrak{R}}_{\lambda} = \|\overset{**}{R}_{ik}(\lambda)\|$ bezeichnen wir als eine Feldresolvente von $\mathfrak{A}$. Wir dürfen nun ebenso weiter schließen, wie auf S. 169 ff., so daß wir uns kürzer fassen können. Nach dem Fundamentalkriterium (§ 48) gibt es eine und nur eine, für $\mu \to -\infty$ verschwindende, von rechts stetige Funktion mit einer Totalschwankung $\leqq 1$ derart, daß

$$\overset{**}{R}_{ik}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\overset{**}{\sigma}_{ik}(\mu)}{\lambda - \mu} \tag{448}$$

gilt. Wir nennen $\|\overset{**}{\sigma}_{ik}(\mu)\|$ eine Feldspektralmatrix von $\mathfrak{A}$; sie wird durch $\{\mathfrak{A}^{\{h_n\}}\}$ „induziert". Die Kopplungsform der Feldspektralmatrix ist auf E durchweg $\geqq 0$ und $\leqq 1$ und nimmt mit μ nicht ab. Man hat ferner $\mathsf{M}(\|\overset{**}{R}_{ik}(\lambda)\|) \leqq \frac{1}{\mathfrak{J}[\lambda]}$. Dies alles bleibt nämlich bei dem Grenzübergang $h_n \to +\infty$ offenbar erhalten.

§ 118. Stabilität der regulären Matrizen.

Die Feldspektralmatrix $\| \overset{**}{\sigma}_{ik}(\mu) \|$ von $\mathfrak{A}$ kann von verschiedenen Umständen abhängen, so von der zugrunde gelegten mittelkonvergenten Folge $\{\mathfrak{A}^{\{n\}}\}$ und außerdem von ihrer induzierenden Teilfolge $\{\mathfrak{A}_i^{\{h_n\}}\}$. Es sei nun verlangt, daß $\mathfrak{A}$ derart ist, daß es dennoch nur eine Feldspektralmatrix, oder was offenbar dasselbe bedeutet, nur eine Feldresolvente hat. In diesem Falle nennen wir $\mathfrak{A}$ statistisch stabil. Wie schwerwiegend diese Stabilitätsannahme ist, geht bereits daraus hervor, daß die Matrizen der mittelkonvergenten Folge nicht statistisch sinnvoll zu sein brauchen, also verschiedene Grenzspektralmatrizen haben können, während wir verlangen, daß der Limes $\overset{**}{\sigma}_{ik}(\mu)$ dennoch eindeutig bestimmt sei. Es liegt also nahe, an Stelle der absoluten Stabilität nur eine *bedingte* zu verlangen, d. h. beim Induzieren der Feldspektralmatrizen für $\mathfrak{A}$ nur virtuelle Störungen einer bestimmten Art zuzulassen, die eben der Natur von $\mathfrak{A}$ oder der Problemestellung angepaßt ist.

Dennoch gibt es statistisch absolut stabile Matrizen: wir beweisen nach Carleman, daß die regulären (also z. B. die beschränkten) Matrizen stabil sind.

Es sei $\mathfrak{c} = (c_1, c_2, \ldots)$ irgendein quadratisch konvergenter Vektor. Da $\overset{*}{\mathfrak{R}}_\lambda^{\{h_n\}}$ eine Grenzresolvente von $\mathfrak{A}^{\{h_n\}}$ ist, so ist, wie wir wissen, (S. 218),

$$x_i^{\{h_n\}} = \sum_{j=1}^{\infty} \overset{*}{R}{}_{ij}^{\{h_n\}}(\lambda)\, c_j \tag{449}$$

gewiß eine quadratisch konvergente Lösung von $(\lambda \mathfrak{E} - \mathfrak{A}^{\{h_n\}})\mathfrak{x} = \mathfrak{c}$, d. h. von

$$\lambda x_i^{\{h_n\}} - \sum_{j=1}^{\infty} a_{ij}^{\{h_n\}} x_j^{\{h_n\}} = c_i. \tag{450}$$

Und zwar gibt es eine nur von λ abhängige Schranke $\varLambda(\lambda)$, derart, daß

$$\sum_{i=1}^{\infty} | x_i^{\{h_n\}} |^2 < \varLambda(\lambda) \cdot \sum_{i=1}^{\infty} | c_j |^2 \tag{451}$$

ausfällt. Dies ist aus (449) unmittelbar abzulesen, wenn man beachtet, daß die Kopplungsform einer jeden Grenzresolvente $\| \overset{*}{R}{}_{ik}^{(n)}(\lambda) \|$ auf **E** absolut $\leqq \dfrac{1}{\mathfrak{J}[\lambda]}$ ist. Dies gilt, wie erwähnt, auch von der Feldresolvente $\| \overset{**}{R}_{ik}(\lambda) \|$, die also ebenfalls beschränkt ist, so daß der Vektor $\mathfrak{x}$ mit

den Komponenten

$$x_i = \sum_{j=1}^{\infty} \overset{**}{R}_{ij}(\lambda)\, c_j \tag{452}$$

eine endliche Norm hat. Da nach Definition $\overset{*}{R}{}_{ik}^{\{h_n\}}(\lambda) \to \overset{**}{R}_{ik}(\lambda)$, $n \to \infty$ gilt, so folgt aus (449), (451), (452) mit Rücksicht auf das Majorantenkriterium der starken Konvergenz die Grenzgleichung $x_i^{\{h_n\}} \to x_i$. Andererseits gilt auch

$$\sum_{j=1}^{\infty} a_{ij}^{\{h_n\}}\, x_j^{\{h_n\}} \to \sum_{j=1}^{\infty} a_{ij}\, x_j, \tag{453}$$

also wegen (450)

$$\lambda x_i - \sum_{j=1}^{\infty} a_{ij}\, x_j = c_i; \qquad (\lambda \mathfrak{E} - \mathfrak{A})\mathfrak{x} = \mathfrak{c}. \tag{454}$$

Um (453) zu beweisen, genügt es, zu zeigen, daß für $n \to +\infty$ beide Differenzen

$$\sum_{j=1}^{\infty} (a_{ij}^{\{h_n\}} - a_{ij})\, x_j^{\{h_n\}}, \qquad \sum_{j=1}^{\infty} a_{ij}(x_j - x_j^{\{h_n\}}) \tag{455}$$

nach Null streben. Für die erste ist dies aber wegen (447), (451) nach der Schwarzschen Ungleichheit trivial, für die zweite folgt es aus dem Majorantenkriterium (§ 57). Also ist (454) richtig. Eine Lösung von (454) wird wegen (452) durch die (gewiß beschränkte) Feldresolvente von $\mathfrak{A}$ geliefert, die irgendeine der Feldresolventen von $\mathfrak{A}$ sein kann. Gibt es also immer nur eine quadratische konvergente Lösung, so kann es nur eine Feldresolvente für $\mathfrak{A}$ geben. M. a. W., jede reguläre Matrix ist statistisch stabil. Dies wollten wir beweisen.

Man hat ferner, analog zu (453), für $n \to \infty$

$$\sum_{j=1}^{\infty} \overset{*}{R}{}_{ij}^{\{h_n\}}(\lambda)\cdot(\delta_{jk}\lambda - a_{jk}^{\{h_n\}}) \to \sum_{j=1}^{\infty} \overset{**}{R}_{ij}(\lambda)\cdot(\delta_{jk}\lambda - a_{jk}). \tag{456}$$

Denn es ist wieder [vgl. (455)]

$$\begin{aligned} &\sum_{j=1}^{\infty} \left\{ \overset{*}{R}{}_{ij}^{\{h_n\}}(\lambda) - \overset{**}{R}_{ij}(\lambda) \right\} (\delta_{jk}\lambda - a_{jk}) \to 0, \\ &\sum_{j=1}^{\infty} \left\{ (\delta_{jk}\lambda - a_{jk}) - (\delta_{jk}\lambda - a_{jk}^{\{h_n\}}) \right\} \overset{*}{R}{}_{ij}^{\{h_n\}}(\lambda) \to 0. \end{aligned} \tag{457}$$

Nun ist der Ausdruck in (456) linkerhand für alle h_n gleich δ_{ik}. Denn die Grenzspektralmatrizen sind gewiß vordere Reziproken (S. 220). Also ist auch der Ausdruck rechterhand in (456) gleich δ_{ik}, d. h. die Feldresolventen sind ebenfalls vordere Reziproken. Sie sind aber auch

hintere Reziproken, da auch die Grenzspektralmatrizen es sind. λ bezeichnet hierbei immer eine beliebige nicht reelle Zahl. Gibt es „Spektrumslücken“, so könnte man auch reelle λ in Betracht ziehen.

Die Sprünge der σ liefern wieder quadratisch konvergente Eigenlösungen. Ist $\mathfrak{A}$ eine statistisch beliebige Q-Matrix, deren sämtliche Potenzen Q-Matrizen sind, so ist (305) nach Carleman für jede Feldspektralmatrix (also nicht nur für eine Grenzspektralmatrix) von $\mathfrak{A}$ richtig. — Dies gibt eine Erklärung für den Sachverhalt, von welchem wir auf S. 242 ausgegangen sind.

Auf eine Umkehrung des vorher bewiesenen Satzes, nach welcher jede „extremale“ quadratisch konvergente Lösung von (454) mittels einer Feldresolvente von $\mathfrak{A}$ in der Gestalt (452) dargestellt werden kann, sowie auf die Frage, inwiefern eine „stabile“ Matrix gewiß regulär ist, kann ich hier leider nicht mehr eingehen.

§ 119. „Statistische“ Hauptachsentransformation.

Es sei $\|\sigma_{ik}(\mu)\|$ irgendeine Feldspektralmatrix von $\mathfrak{A}$. Wir erklären das Punktspektrum usf. von $\mathfrak{A}$ wie auf S. 161 bzw. S. 177. Doch ist jetzt zu beachten, daß es z. B. denkbar ist, daß diese Punktmengen jetzt vielleicht nicht mehr eindeutig durch $\mathfrak{A}$ bestimmt sind, indem sie auch von der zugrunde gelegten Grenzspektralmatrix abhängen könnten. Wegen der ersten Relation (422), die auch für Feldspektralmatrizen gilt, hat jedes Spektrum mindestens einen Punkt, ist also nie leer. Um feste Vorstellungen zu haben, wollen wir annehmen, daß $\mathfrak{A}$ regulär ist. Dann gibt es gewiß nur ein Punktspektrum, nur ein Streckenspektrum und nur ein Häufungsspektrum, da es auch nur eine Feldspektralmatrix gibt, die übrigens nach Carleman unseren infinitesimalen Orthogonalitätsrelationen gewiß genügt; siehe S. 225. Ferner ist jede Spektralmatrix jeder Matrix $\mathfrak{A}$ beschränkt (S. 244), und auch jede unitäre Matrix $\mathfrak{U}$ ist beschränkt. Es liegt also nahe, daß Hauptachsenproblem von $\mathfrak{A}$ wieder „hinter das Integralzeichen“ zu verlegen, also die Aufgabe rein statistisch zu fassen und sich mit einer unitären Matrix $\mathfrak{U}$ begnügen, bei welcher die Matrix $\mathfrak{U}\,\|\sigma_{ik}(\mu)\|\,\mathfrak{U}^{-1}$, wie bei (262), in einen rein streckenspektralen Bestandteil und außerdem in eine Summe von (diagonalen) punktspektralen Bestandteilen zerfällt, die von jenem und voneinander getrennt sind.

Wir haben bei den beschränkten Matrizen diesen „statistischen“ Weg nur aus methodischen Gründen verfolgt. In das Schlußresultat der Hauptachsentransformation, daß nämlich $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ in eine Summe

getrennter Diagonalmatrizen und einer von diesen getrennten, diagonal irreduziblen (d. h. rein streckenspektralen) Matrix $\|\tilde{a}_{ik}\|$ zerfällt, geht ja die Spektralmatrix überhaupt nicht ein, so daß wir auch direkt an $\mathfrak{A}$ hätten anknüpfen können. Ist aber $\mathfrak{A}$ eine nicht beschränkte Q-Matrix, so stößt dies auf Schwierigkeiten. Vielleicht ist nämlich $\mathfrak{A}$ derart, daß z. B. $\mathfrak{U}\mathfrak{A}$ mit $\mathfrak{U}^{-1}$ überhaupt nicht multipliziert werden kann, indem die Reihe, welche das (i, k)-te Element der Matrix $(\mathfrak{U}\mathfrak{A})\mathfrak{U}^{-1}$ bilden würde, divergiert. Es scheint also zunächst angemessen zu sein, sich auf die Betrachtung der immer beschränkten Spektralmatrizen zu beschränken. Ist $\mathfrak{A}$ beschränkt, so transformieren sich $\mathfrak{A}$ und $\|\sigma_{ik}(\mu)\|$ kogredient. Insbesondere war das Spektrum unitär invariant. Jetzt entsteht aber eine Schwierigkeit. Gesetzt, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ bei der nicht beschränkten Matrix $\mathfrak{A}$ sich als $(\mathfrak{U}\mathfrak{A})\mathfrak{U}^{-1}$ oder als $\mathfrak{U}(\mathfrak{A}\mathfrak{U}^{-1})$ bilden läßt. Dann leuchtet es überhaupt nicht ein, daß z. B. die Grenzspektralmatrix der etwa statistisch sinnvollen Matrix $\mathfrak{A}$ durch $\mathfrak{U}$ in die Grenzspektralmatrix von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ übergehen muß. Denn ein Eigenwert von $\mathfrak{A}_{[n]}$ ist eben nur n-ären unitären Transformationen gegenüber invariant. Anders ausgedrückt: die Gleichung $\mathfrak{U}_{[n]}\mathfrak{A}_{[n]}\mathfrak{U}^*_{[n]} = (\mathfrak{U}\mathfrak{A}\mathfrak{U}^*)_{[n]}$ ist im allgemeinen unrichtig — auch wenn $\mathfrak{A}$ und $\mathfrak{U}$ durchweg finit sind — und die Abschnitte der für den Hilbertschen Raum unitären Matrix $\mathfrak{U}$ sind, mit Ausnahme von leicht diskutierbaren Grenzfällen, keine unitäre Matrizen für die bez. Abschnittsräume, so daß E durch $\mathfrak{U}$ eineindeutig in sich übergehen kann, ohne daß dies für $\mathsf{E}_{[n]}$, $\mathfrak{U}_{[n]}$ gelten würde; vgl. übrigens die Fußnote auf S. 66. Außerdem steht es nicht fest, daß $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$ statistisch sinnvoll ist, wenn $\mathfrak{A}$ es war. Nach einer freundlichen mündlichen Mitteilung von Herrn J. v. Neumann kann man z. B. jede reguläre nicht beschränkte Matrix in eine nicht reguläre unitär überführen, man kann eine Matrix von durchweg positivem Spektrum in eine solche mit durchweg negativem Spektrum transformieren, usf.

Doch wird man bei „vernünftigen“ $\mathfrak{A}$ wohl immer nur passend gewählte „vernünftige“ $\mathfrak{U}$ betrachten müssen. Z. B. die Formeln (328), (329) des Korrespondenzprinzips stehen auch bei nicht beschränkten Matrizen zur Verfügung. Überhaupt könnte man das Korrespondenzprinzip bei gehöriger Vorsicht auf nicht beschränkte Matrizen unmittelbar übertragen. Entsprechendes gilt erst recht für die multiplikativen Gruppenklassen. Für die Quantenmechanik würde außerdem die Transformationstheorie der finiten $\mathfrak{A}$ mittels finiter $\mathfrak{U}$ sehr wesentlich in Betracht kommen, die Toeplitz gelegentlich angedeutet, jedoch leider nicht veröffentlicht hat. Aus den v. Neumannschen „Paradoxien“ geht hervor, wie tiefliegend solche Sätze sein müssen, und wie wenig selbstverständlich es ist, daß dergleichen bei beschränkten Matrizen gewiß nicht vorkommen kann.

§ 120. Über die Stabilität des reduzierten Spektrums.

Es sei $\mathfrak{A}$ eine reguläre Matrix, $\|\sigma_{ik}(\mu)\|$ ihre (völlig bestimmte) Spektralmatrix. Wir haben in § 116 die Matrix $\mathfrak{A}$ einer virtuellen Störung ausgesetzt, indem wir außer $\mathfrak{A}$ noch

$$\mathfrak{A} + \delta\mathfrak{A}, \qquad \delta\mathfrak{A} = (\mathfrak{A}^{\{n\}} - \mathfrak{A}) \{\rightarrow\} \mathfrak{O} = \|0\|$$

betrachtet haben. Da $\mathfrak{A}$ regulär, also statistisch stabil ist, so war dabei $d\sigma_{ik}^{\{n\}} \rightarrow d\sigma_{ik}$. Wir wollen hier ein etwas anders gerichtetes Ergebnis von H. Weyl ohne Beweis mitteilen. Es möge sich nunmehr nicht um virtuelle, sondern endliche (d. h. feste), jedoch „kleine“ Störungen handeln. Wir fragen, wie das Spektrum auf solche Störungen reagiert. Als eine „kleine“ Störung von $\mathfrak{A}$ bezeichnen wir den Ersatz von $\mathfrak{A}$ durch $\mathfrak{A} + \mathfrak{B}$, wobei $\mathfrak{B}$ eine feste *vollstetige* Hermitesche Matrix bezeichnet. Vollstetig heißt die Matrix $\mathfrak{B}$, wenn sie beschränkt ist und durch eine unitäre Substitution $\mathfrak{U}$ auf die Gestalt $\mathfrak{U}\mathfrak{B}\mathfrak{U}^{-1} = \|\lambda_i \delta_{ik}\|$, $\lambda_i \rightarrow 0$ gebracht werden kann. Hierfür ist, wie wir wissen, notwendig und hinreichend, daß es kein Streckenspektrum gibt und daß alle von Null verschiedenen Eigenwerte von endlichem Gewicht sind und sich höchstens im Nullpunkte häufen. N. B., $\mathfrak{B} = \|v_{ik}\|$ selbst braucht nicht von der Gestalt $\|\lambda_i \delta_{ik}\|$ zu sein. Eine hinreichende Bedingung für die Vollstetigkeit von $\mathfrak{B}$ besteht in der Konvergenz der Doppelreihe $\sum_{i=1}^{\infty}\sum_{i=1}^{\infty} |v_{ik}|^2$, d. i. der „Normalspur“ von S. 28. Notwendig ist diese Bedingung bereits darum nicht, weil $\|v_{ik}\|$ vollstetig sein kann, ohne daß die Matrix $\| |v_{ik}| \|$ der Beträge vollstetig, ja überhaupt beschränkt wäre. — Notwendig und hinreichend ist für die Vollstetigkeit der Hermiteschen Matrix $\mathfrak{B}$, daß ihre Kopplungsform auf der Hilbertschen Einheitskugel gleichmäßig konvergiert. Das Produkt einer vollstetigen Matrix in eine beschränkte Matrix ist stets vollstetig. Da die Einheitsmatrix offenbar nicht vollstetig ist, so kann daher keine vollstetige Matrix eine beschränkte Reziproke haben. — Wir haben diese Sätze ohne Beweis angegeben. Sie rühren von Hilbert her, der die Theorie der vollstetigen Matrizen voll entwickelt hat (und zwar auch für nicht Hermitesche Matrizen; die Vollstetigkeit ist im allgemeinen Fall durch die Gleichmäßigkeit der Konvergenz der Kopplungsform zu definieren).

Es sei $\mathfrak{A}$ eine reguläre Matrix, $\mathfrak{s}$ das Spektrum, $\mathfrak{b}$ das Streckenspektrum („Bandenspektrum“), $\mathfrak{p}$ das Punktspektrum und $\mathfrak{h}$ das Häufungsspektrum. Es ist wesentlich, daß $\mathfrak{h}$ genau in dem auf S. 161 erklärten Sinne verstanden werden soll. Wir führen noch eine (schlichte) Punktmenge ein, die wir mit $\mathfrak{r}$ bezeichnen. Sie enthalte die folgen-

den Punkte und nur diese: 1. alle Punkte von $\mathfrak{b}$; 2. alle Punkte von $\mathfrak{h}$; 3. diejenigen Punkte von $\mathfrak{p}$, die Eigenwerte von unendlich hohem Gewicht darstellen (so daß die zu $\mu\mathfrak{E} - \mathfrak{A}$ gehörigen homogenen Gleichungen unendlich viele linear unabhängige quadratisch konvergente Lösungen haben). Die dadurch festgelegte Punktmenge $\mathfrak{r}$ nennen wir das reduzierte Spektrum von $\mathfrak{A}$. Vollstetig ist die (beschränkte) Matrix offenbar dann und nur dann, wenn das reduzierte Spektrum auf den Nullpunkt der Spektrumsgerade zusammenschrumpft. Allgemeiner ist das reduzierte Spektrum $\mathfrak{r} = \mathfrak{s}' + \mathfrak{p}_\infty$, wobei $\mathfrak{s}'$ die Gesamtheit der Häufungsstellen des schlicht gedachten Gesamtspektrums $\mathfrak{s}$, und $\mathfrak{p}_\infty$ die Vereinigung derjenigen Eigenwerte unendlich hoher Vielfachheit bezeichnet, die $\mathfrak{s}'$ noch nicht angehören. Offenbar ist $\mathfrak{r}$ immer abgeschlossen.

Weyl hat nun bewiesen, daß das reduzierte Spektrum jeder beschränkten Hermiteschen Matrix beliebigen „kleinen" Störungen gegenüber unempfindlich ist: $\mathfrak{A}$ und $\mathfrak{A} + \mathfrak{B}$ haben dasselbe reduzierte Spektrum, wenn $\mathfrak{B}$ vollstetig ist. Wie Carleman gezeigt hat, bleibt dies richtig, auch wenn $\mathfrak{A}$ nur regulär ist. Weyl hat ferner mittels der Haarschen Orthogonalsysteme bewiesen, daß das reduzierte Spektrum gewissermaßen die einzige Invariante bei kleinen Spektrumsstörungen ist. Hat z. B. $\mathfrak{A}$ nur ein Streckenspektrum, das etwa das Einheitsintervall sein möge, so kann man eine vollstetige Matrix $\mathfrak{B}$ derart angeben, daß $\mathfrak{A} + \mathfrak{B}$ kein Streckenspektrum mehr besitzt, also auf die Diagonalform gebracht werden kann. Da jedoch das Einheitsintervall, als Streckenspektrum, dem reduzierten Spektrum angehört, so müssen die „diskreten" Eigenwerte von $\mathfrak{A} + \mathfrak{B}$ (d. h. sein abzählbares Punktspektrum) auf dem Einheitsintervall überall dicht liegen.

Über die für die Quantenmechanik so wichtige Frage einer tatsächlichen Behandlung von Spektrumsstörungen liegen in der mathematischen Literatur leider keine befriedigenden Ergebnisse vor. Insbesondere steht eine sinngemäße Übertragung der Hill-Brownschen Methoden der Mondtheorie noch aus.

Anhang.

Skizze einer Spektraltheorie der fastperiodischen Funktionen.

Zur Illustration der Spektraltheorie soll eine Klasse von Matrizen betrachtet werden, welche die Mechanik der klassischen Schwingungen in die Matrizensprache übersetzt. Es handelt sich um eine sinngemäße Übertragung der Toeplitzschen Theorie der Laurentmatrizen auf die fastperiodischen Funktionen von H. Bohr.

Wir wollen zunächst den Hauptsatz über diese Funktionen vorausschicken. Die für $-\infty < t < +\infty$ erklärte, stetige, nicht notwendig reellwertige Funktion $f(t)$ möge wie folgt beschaffen sein. Ist $\varepsilon > 0$ fest gewählt, so gibt es eine Folge $\tau_\nu = \tau_\nu(\varepsilon)$; $\nu = 1, 2, \ldots$ von „Verschiebungszahlen" derart, daß einerseits $|f(t + \tau_\nu(\varepsilon)) - f(t)| \leqq \varepsilon$ für alle t und alle ν gilt, und daß andererseits die Folge ins Unendliche wächst, ohne dabei beliebig große Lücken aufzuweisen:

$$\tau_1(\varepsilon) < \tau_2(\varepsilon) < \cdots \to +\infty; \qquad \limsup_{\nu \to +\infty} (\tau_{\nu+1}(\varepsilon) - \tau_\nu(\varepsilon)) < +\infty;$$

hierbei darf der Limes superior mit ε^{-1} unendlich werden, doch sollen die aufgezählten Bedingungen bei beliebig kleinem (festen) ε erfüllt sein. Die Funktion $f(t)$ heißt dann fastperiodisch.

Eine periodische Funktion ist dann und nur dann fastperiodisch, wenn sie stetig ist. Jede fastperiodische Funktion ist für $-\infty < t < +\infty$ gleichmäßig stetig, also insbesondere beschränkt. Konvergiert eine Folge von fastperiodischen Funktionen $f_n(t)$ *für alle t gleichmäßig* gegen eine Grenzfunktion $f(t)$, so ist auch $f(t)$ fastperiodisch. Die gleichmäßige Konvergenz ist also in dem Sinne $\mathfrak{m}(f - f_n) \to 0$ zu verstehen, wobei $\mathfrak{m}(g)$ die obere Grenze von $|g(t)|$ für $-\infty < t < +\infty$ bedeutet. Die untere Grenze von $|g(t)|$ möge entsprechend mit $\mathfrak{n}(g)$ bezeichnet werden. Es kann g fastperiodisch, durchweg > 0 und dennoch derart sein, daß $\mathfrak{n}(g) = 0$ ausfällt, weil $g(t)$ der Null beliebig

nahe kommt (ohne zu verschwinden; $g(t)$ ist zwar stetig, das Intervall ist indessen unendlich). Eine ganze rationale Funktion von endlich vielen fastperiodischen Funktionen (m. a. W., Summe und Produkt zweier fastperiodischen Funktionen) ist wieder fastperiodisch, also ist es auch jede endliche trigonometrische Summe

$$(1) \qquad c_1 e_{\xi_1}(t) + c_2 e_{\xi_2}(t) + \cdots + c_m e_{\xi_m}(t),$$

wobei c_k konstant und

$$(2) \qquad e_\xi(t) = e^{\xi t \sqrt{-1}}; \qquad \sqrt{-1} = i; \qquad \xi \gtreqless 0$$

ist; denn die stetige, periodische Funktion $e_\xi(t)$ ist gewiß fastperiodisch. Mithin ist eine Funktion gewiß fastperiodisch, wenn sie durch eine Folge $\{f_n\}$ gleichmäßig approximiert werden kann, wobei jedes f_n von der Gestalt (1) ist (die ξ dürfen dabei von n abhängen und m darf mit n über alle Grenzen wachsen). Es gilt nun — in Verallgemeinerung des Weierstraßschen Satzes über stetige, periodische Funktionen — die Umkehrung, daß zu jeder fastperiodischen Funktion $f(t)$ eine Folge $\{f_n\}$ von endlichen trigonometrischen Summen angegeben werden kann derart, daß $\mathfrak{m}(f - f_n) \to 0$, $n \to +\infty$ gilt. Dies ist der Hauptsatz der Bohrschen Theorie.

Wir wollen hier sogleich einen für das Spätere wichtigen Begriff einführen. Bei jeder fastperiodischen Funktion ist nach Bohr der Grenzmittelwert

$$(3) \qquad \mathfrak{M}(f) = \lim_{T \to +\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\, dt$$

vorhanden (doch ist dies nur eine notwendige Bedingung für den fastperiodischen Charakter). Bezeichnet α irgendeine Konstante und $\mathfrak{k}_T(\alpha)$ diejenige Teilmenge des Intervalles $-T \leqq t \leqq T$, auf welcher $f(t) = \alpha$ ausfällt, so ist unter Umständen auch der Grenzwert

$$(4) \qquad \mathfrak{N}_\alpha(f) = \lim_{T \to +\infty} \frac{1}{2T} \int_{\mathfrak{k}_T(\alpha)} dt$$

vorhanden. Bei derartigen Betrachtungen haben die folgenden beiden, aus f abgeleiteten Scharen von stetigen Funktionen eine Rolle:

$$f_\alpha^+(t) = f(t) \quad \text{für} \quad f(t) \leqq \alpha; \qquad f_\alpha^+(t) = \alpha \quad \text{für} \quad f(t) \geqq \alpha$$

und

$$f_\alpha^-(t) = f(t) \quad \text{für} \quad f(t) \geqq \alpha; \qquad f_\alpha^-(t) = \alpha \quad \text{für} \quad f(t) \leqq \alpha.$$

Aus der ε-Definition der Fastperiodizität geht unmittelbar hervor, daß die beiden Funktionen $f_\alpha^{\pm}(t)$ fastperiodisch sind, wenn f es ist. „Im allgemeinen" wird $\mathfrak{N}_\alpha(f)$ für alle α verschwinden. Doch kann $\mathfrak{N}_\alpha(f)$ auch $\neq 0$ sein. Es sei z. B. f stetig, periodisch und derart, daß es in einem ganzen Intervall den Wert α hat. Allgemein bezeichnet $\mathfrak{N}_\alpha(f)$ die Wahrscheinlichkeit a priori (relative Häufigkeit) des Wertes α in dem Wertevorrate von $f(t)$. Für (4) kann man nämlich offenbar auch

$$(5)\qquad \mathfrak{N}_\alpha(f) = \lim_{T\to+\infty} \frac{\operatorname{mes} \mathfrak{f}_T(\alpha)}{2T}$$

schreiben, wobei $\operatorname{mes} \mathfrak{a}$ das lineare Maß (mesure, Inhalt) einer Punktmenge $\mathfrak{a}$ bezeichnet.

Wir werden später nicht diese, sondern die „integralen" Wahrscheinlichkeiten benötigen. Es sei die fastperiodische Funktion f zunächst reellwertig, und es bezeichne $\mathfrak{t}_T(\alpha)$ diejenige Teilmenge des Intervalles $-T \leqq t \leqq T$, auf welcher $f(t) \leqq \alpha$. Dann läßt sich, in Verallgemeinerung von (3), auch die Existenz des Grenzwertes

$$(6)\qquad \mathfrak{M}_\alpha(f) = \lim_{T\to+\infty} \frac{1}{2T} \int\limits_{\mathfrak{t}_T(\alpha)} f(t)\,dt$$

„im wesentlichen", d. h. in dem auf S. 77 präzisierten Sinne gewiß behaupten, und in demselben Sinne auch die Existenz von $\mathfrak{M}_0(f-\alpha)$. Hierbei ist beidemal α die unabhängige Veränderliche der Grenzfunktion. — Auf eine einfache Herleitung dieser Behauptungen, sofern sie für unsere Zwecke in Betracht kommen, kommen wir auf S. 270 zurück. Ein Beweis, der auch etwa unter Benutzung der unstetigen Verallgemeinerungen der fastperiodischen Funktionen geführt wird, würde zu weit führen und soll hier übergangen werden.

Es bezeichne $\boldsymbol{n}(f)$ die untere, $\boldsymbol{m}(f)$ die obere Grenze von f für alle t. Diese Zahlen sind also, im Gegensatz zu den Zahlen $\mathfrak{n}(f)$, $\mathfrak{m}(f)$, evtl. negativ und nur für reelle f definiert. Es bezeichne $\boldsymbol{n}_T(f)$ das Minimum, $\boldsymbol{m}_T(f)$ das Maximum von $f(t)$ auf dem Intervall $-T \leqq t \leqq T$, so daß

$$(7)\qquad \boldsymbol{n}(f) \leqq \boldsymbol{n}_T(f), \qquad \boldsymbol{m}_T(f) \leqq \boldsymbol{m}(f)$$

gilt, wobei das Zeichen $\leqq$ für $T \to +\infty$ gewiß in $=$ übergeht. Man erkläre in dem Intervall

$$(8)\qquad \boldsymbol{n}_T(f) \leqq \varphi \leqq \boldsymbol{m}_T(f),$$

wobei T fest gewählt ist, eine Funktion $\nu_T(\varphi)$ wie folgt:

$$\nu_T(\varphi) = \frac{\operatorname{mes} \mathfrak{t}_T(\varphi)}{2T}, \tag{9}$$

so daß $\nu_T(\varphi)$ die Wahrscheinlichkeit a priori dafür ist, daß die Funktion f auf dem Intervall $-T \leqq t \leqq T$ nicht größer als φ ausfällt. Der Grenzwert

$$\nu(\varphi) = \lim_{T \to +\infty} \nu_T(\varphi) \tag{10}$$

stellt daher, sofern er existiert, die durchschnittliche relative Verweilzeit der Kurve $f = f(t)$ in der Halbebene $f \leqq \varphi$ oder in dem Streifen $n(f) \leqq f \leqq \varphi$ der (t, f)-Ebene dar. Nun ist aber der Grenzwert (*10*) „im wesentlichen", d. h. an allen Stetigkeitsstellen der Grenzfunktion stets vorhanden. Beim Beweis kann ohne Einschränkung der Allgemeinheit angenommen werden, daß $\varphi \neq 0$ ist. Denn der Streifen darf in sich verschoben werden. Wir fassen nun φ als eine Konstante auf und wenden die vorher erwähnten Sätze auf die fastperiodische Funktion $f(t) - \varphi$ an, indem wir $\alpha = 0$ setzen. So folgt die Existenz des Grenzwertes

$$\mathfrak{M}_0(f - \varphi) = \lim_{T \to +\infty} \frac{1}{2T} \int (f - \varphi)\, dt, \tag{11}$$

wobei die Integration sich auf die Menge derjenigen t erstreckt, für welche $-T \leqq t \leqq T$ und $f - \varphi \leqq 0$, d. h. $f \leqq \varphi$ ist, also auf die (zu f gehörige) Menge $\mathfrak{t}_T(\varphi)$. Da der Grenzwert

$$\lim_{T \to +\infty} \frac{1}{2T} \int\limits_{\mathfrak{t}_T(\varphi)} f(t)\, dt = \mathfrak{M}_\varphi(f) \tag{12}$$

wegen (*6*) vorhanden ist, so folgt die Existenz des Grenzwertes

$$\lim_{T \to +\infty} \frac{1}{2T} \int\limits_{\mathfrak{t}_T(\varphi)} \varphi\, dt = \mathfrak{M}_\varphi(f) - \mathfrak{M}_0(f - \varphi). \tag{13}$$

Wegen

$$\frac{1}{2T} \int\limits_{\mathfrak{t}_T(\varphi)} \varphi\, dt = \frac{\varphi}{2T} \int\limits_{\mathfrak{t}_T(\varphi)} dt = \varphi \frac{\operatorname{mes} \mathfrak{t}_T(\varphi)}{2T} = \varphi\, \nu_T(\varphi) \tag{14}$$

folgt daraus die Existenz von (*10*), da, wie erwähnt, vorausgesetzt werden darf, daß $\varphi \neq 0$ ausfällt. Die Aussage, „der Grenzwert existiert", ist dabei in dem Sinne [$\to$] der Konvergenz im wesentlichen zu ver-

stehen. — Für das folgende kommt unter den erwähnten Grenzwertsätzen nur die Existenz von (*10*), d. h.

$$\nu_T(\varphi)\,[\rightarrow]\,\nu(\varphi), \qquad T \rightarrow +\infty \tag{15}$$

in Betracht. Auf einen, durch die matrizentheoretischen Zusammenhänge nahegelegten höchst einfachen Beweis von (*15*) kommen wir auf S. 270 zurück.

Die Definition (*9*) der Funktion $\nu_T(\varphi)$ wollten wir nur auf dem Intervall (*8*) in Anspruch nehmen. Da im allgemeinen kein (endliches) T mit $m_T(f) = m(f)$ oder $n_T(f) = n(f)$ existiert, indem die monotonen Funktionen m_T, n_T ihren Grenzwert m, n evtl. nur nach Vornahme des Grenzüberganges $T \rightarrow +\infty$ erreichen, so wollen wir (*15*) bei jedem f nur für

$$n(f) < \varphi < m(f) \tag{16}$$

als eine Definition auffassen, so daß $\nu(f)$ in den beiden Endpunkten von (*16*) zunächst noch nicht erklärt ist. An den eventuellen Unstetigkeitsstellen der Grenzfunktion $\nu(\varphi)$, wo wir die Existenz des Grenzwertes nicht behauptet haben, möge $\nu(\varphi)$ durch die Forderung der rechtsseitigen Stetigkeit festgelegt sein. Dann ist $\nu(\varphi)$ auf (*16*) bei jedem f eindeutig erklärt. Von dem trivialen Fall $f \equiv$ const wird dabei abgesehen.

Wir erweitern die Funktion $\nu_T(\varphi)$, indem wir sie auch für $n(f) \leqq \varphi \leqq n_T(f)$ und $m_T(f) \leqq \varphi \leqq m(\varphi)$ erklären: es sei $\nu_T(\varphi) \equiv \nu_T(n_T(f))$ auf dem unteren, $\equiv \nu_T(m_T(f))$ auf dem oberen Zusatzintervall. Die Funktion $\nu_T(\varphi)$ ist jetzt also für

$$n(f) \leqq \varphi \leqq m(f) \tag{17}$$

erklärt, und für $\varphi \leqq n_T(f)$ und für $\varphi \geqq m_T(f)$ konstant. Sie ist offenbar durchweg $\geqq 0$, nicht abnehmend, und für $\varphi \geqq m_T(f)$ gleich 1. Hingegen braucht sie für $\varphi \leqq n_T(f)$ nicht zu verschwinden (man denke nur an eine stetige, periodische Funktion, die ihr Minimum längs gewissen Strecken annimmt). Wir legen den Wert von $\nu(\varphi)$ im oberen Endpunkte $\varphi = m(f)$ von (*16*) durch die Definition $\nu(m(f)) = 1$ fest. Dann gilt (*10*) auch für $\varphi = m(f)$. Wir wollen endlich sinngemäß [vgl. (*5*)] $\nu(n(f)) = \mathfrak{R}_{n(f)}(f)$ setzen, sofern dieser Grenzwert existiert.

Man hat

$$\frac{1}{2T}\int_{-T}^{T} f(t)\,dt = \int_{n_T(f)}^{m_T(f)} \varphi\, d\nu_T(\varphi). \tag{18}$$

Denn die Stieltjesschen Näherungssummen des zweiten Integrals sind, wie aus der Definition von $\nu_T(\varphi)$ hervorgeht, eben die Lebesgueschen Näherungssummen des ersten. Der mit dem Lebesgueschen Integralbegriff nicht vertraute Leser möge das linkerhand stehende Integral, das freilich auch als ein Riemannsches Integral aufgefaßt werden kann (f ist ja stetig), als Flächeninhalt deuten und sich die Richtigkeit der Formel etwa für stückweise monotone Funktionen f klarmachen. — Da $\nu_T(\varphi)$ für $\varphi \leqq \boldsymbol{n}_T(f)$ und $\varphi \geqq \boldsymbol{m}_T(\varphi)$ konstant ist, so kann man für (*18*) auch

$$(19)\qquad \frac{1}{2T}\int_{-T}^{T} f(t)\,dt = \int_{\boldsymbol{n}(f)}^{\boldsymbol{m}(f)} \varphi\, d\nu_T(\varphi)$$

schreiben. Es kann dabei auch $\boldsymbol{n}_T(f) = \boldsymbol{n}(f)$ eintreten, und dann ist evtl. $\nu_T(\boldsymbol{n}(f)+0) \neq \nu_T(\boldsymbol{n}(f))$. Für alle Fälle ist

$$(20)\qquad \frac{1}{2T}\int_{-T}^{T} f(t)\,dt = J_1^T + J_2^T;$$

$$J_1^T = \int_{\boldsymbol{n}(f)+0}^{\boldsymbol{m}(f)} \varphi\, d\nu_T(\varphi), \quad J_2^T = \boldsymbol{n}(f)\,[\nu_T(\boldsymbol{n}_T(f)+0) - \nu_T(\boldsymbol{n}_T(f))].$$

Nun gilt (*15*) auf (*16*), und (*10*) kann in dem oberen Endpunkte des Intervalles (*16*) gewiß behauptet werden. Ferner sind die Funktionen $\nu_T(\varphi)$ nicht negativ, nicht abnehmend und $\leqq 1$, so daß aus dem Hellyschen Satz über gliedweise Integration offenbar

$$(21)\qquad \lim_{T\to+\infty} J_1^T = \int_{\boldsymbol{n}(f)+0}^{\boldsymbol{m}(f)} \varphi\, d\nu(\varphi)$$

folgt. Die linke Seite von (*20*) strebt für $t \to +\infty$ der Grenze $\mathfrak{M}(f)$ zu. Mithin existiert auch $\lim J_2^T$, also auch der Grenzwert

$$(22)\qquad C = \lim_{T\to+\infty} [\nu_T(\boldsymbol{n}(f)+0) - \nu_T(\boldsymbol{n}(f))];$$

der Fall $\boldsymbol{n}(f) = 0$ kann nämlich ausgeschlossen werden, vgl. S. 254.

Definieren wir nun die „Dichtefunktion von f" wie folgt[1]):

$$(23)\qquad \begin{aligned} \omega(\varphi) &= \nu(\varphi) && \text{für } \boldsymbol{n}(f) < \varphi \leqq \boldsymbol{m}(f),\\ &= \nu(\boldsymbol{n}(f)+0) - C && \text{für } \varphi = \boldsymbol{n}(f),\\ &= \omega(\boldsymbol{m}(f)) = 1 && \text{für } \varphi > \boldsymbol{m}(f),\\ &= \omega(\boldsymbol{n}(f)) && \text{für } \varphi < \boldsymbol{n}(f), \end{aligned}$$

---

[1]) Die Wahrscheinlichkeitsdichte ist freilich nicht $\omega(\varphi)$, sondern $d\omega(\varphi) : d\varphi$.

so ist wegen (*22*) und (*20*)

$$(24)\qquad \int\limits_{\boldsymbol{n}(f)}^{\boldsymbol{n}(f)+0} \varphi\, d\omega(\varphi) = \boldsymbol{n}(f)\left[\omega(\boldsymbol{n}(f)+0) - \omega(\boldsymbol{n}(f))\right] = \boldsymbol{n}(f)\, C = \lim_{T\to+\infty} J_2^T,$$

und in (*21*) darf ω an Stelle von ν geschrieben werden, so daß mit Rücksicht auf (*20*)

$$(25)\qquad \mathfrak{M}(f) = \int\limits_{\boldsymbol{n}(f)}^{\boldsymbol{m}(f)} \varphi\, d\omega(\varphi) = \int\limits_{-\infty}^{+\infty} \varphi\, d\omega(\varphi)$$

gilt.

Die Funktion $\omega(\varphi)$ ist auf (*16*) gewiß nicht abnehmend. Es wird sich aus dem matrizentheoretischen Zusammenhang indirekt herausstellen, daß $\omega(\varphi)$ auch auf (*17*) [also überhaupt] nicht abnehmend und von der Totalschwankung Eins sein muß[1]), was freilich auch direkt eingesehen werden kann. Da demnach

$$(26)\qquad \int\limits_{\boldsymbol{n}(f)}^{\boldsymbol{m}(f)} d\omega(\varphi) = \omega(\boldsymbol{m}(f)) - \omega(\boldsymbol{n}(f)) = 1$$

gilt, so folgt daraus, daß $\omega(\boldsymbol{n}(f)) = 0$ ist. Denn nach (*23*) ist $\omega(\boldsymbol{m}(f)) = 1$. Aus $\omega(\boldsymbol{n}(f)) = 0$ folgt übrigens nach (*23*), daß der Grenzwert (*22*) gleich $\nu(\boldsymbol{n}(f)+0)$ sein muß. Da $\omega(\boldsymbol{n}(f)) = 0$ immer gilt, während $\nu(\boldsymbol{n}(f))$, wie erwähnt (S. 253), auch > 0 sein kann, so ist die Dichtefunktion $\omega(\varphi)$ von der Verteilungsfunktion $\nu(\varphi)$ im allgemeinen wohl zu unterscheiden, sobald es sich nicht nur um (*16*) handelt.

Die Funktion $\omega(\varphi)$ erleidet für $\varphi = \varphi_0$ gewiß einen Sprung, wenn $\mathfrak{N}_{\varphi_0}(f) \neq 0$ ist. Doch können solche φ_0, die auf (*17*) auch überall dicht vorkommen können, nur auf (*17*) liegen, da $\omega(\varphi)$ außerhalb (*17*) konstant, also erst recht stetig ist. Es gilt endlich der für unsere Zwecke wichtige Satz, daß $\omega(\varphi)$ auf (*17*) keine Konstanzstrecke haben kann: aus

$$(27)\qquad \boldsymbol{n}(f) \leqq \alpha < \beta \leqq \boldsymbol{m}(f)$$

folgt nicht nur $\omega(\alpha) \leqq \omega(\beta)$, sondern auch $\omega(\alpha) < \omega(\beta)$.

Der Beweis ergibt sich aus dem folgenden Satze von Bohr: Ist $f(t)$ reellwertig, nichtnegativ, *fastperiodisch* und derart, daß $\mathfrak{M}(f) = 0$ ist, so verschwindet $f(t)$ identisch. Der Satz gilt (auch für analytische

---

[1]) Dies besagt nur, daß die Wahrscheinlichkeit a priori dafür, daß die Kurve $f = f(t)$ auf der Ebene (t, f) liegt, gleich Eins ist.

Funktionen) nicht mehr, wenn die kursiv gesetzte Bedingung fallengelassen wird: der Operator $\mathfrak{M}$ ist eben kein Integraloperator, sondern eine Mittelbildung, und eine fastperiodische Wiederkehr ist für die Richtigkeit des Satzes wesentlich. — Setzt man den Satz, daß $\omega(\varphi)$ nicht abnimmt und die Totalschwankung Eins hat, als bekannt voraus, so kann der Bohrsche Satz aus (23), (26) unmittelbar abgelesen werden.

Nehmen wir nun zwecks Zurückführung ad absurdum an, daß $\omega(\alpha) = \omega(\beta)$ ausfällt, trotzdem (27) vorausgesetzt ist. Da $\omega(\varphi)$ nicht abnimmt, so ist dann $\omega(\varphi) \equiv \omega(\alpha)$ für $\alpha < \varphi < \beta$. Wir wählen zwei Zahlen α', β' derart, daß $\alpha < \alpha' < \beta' < \beta$ gilt und daß α' und β' dasselbe Vorzeichen haben; sie seien etwa positiv. Es bezeichne $g(t)$ die Funktion, die gleich

$$(28) \qquad \alpha', \qquad f(t), \qquad \beta' \, (> \alpha' > 0)$$

ist, je nachdem

$$(29) \qquad f(t) \leqq \alpha', \qquad \alpha' \leqq f(t) \leqq \beta', \qquad f(t) \leqq \beta'$$

gilt. Die Funktion $g(t)$ ist gewiß fastperiodisch, und zwar aus demselben Grunde wie $f_\varphi^{\pm}(t)$ auf S. 253. Bezeichnet $\varrho(\varphi)$ die Dichtefunktion von g, so gilt offenbar $\varrho(\varphi) \equiv \omega(\varphi) = \omega(\alpha)$ für $\alpha' \leqq \varphi \leqq \beta'$. Andererseits ist $\boldsymbol{n}(g) = \alpha'$, $\boldsymbol{m}(g) = \beta'$, also muß $\varrho(\varphi) \equiv \omega(\alpha')$ für $\varphi \leqq \alpha'$ und $\varrho(\varphi) \equiv \omega(\beta')$ für $\varphi \geqq \beta'$ und daher $\varrho(\varphi) \equiv \omega(\alpha) =$ Konstante für $-\infty < \varphi < +\infty$ gelten, so daß $\int\limits_{-\infty}^{+\infty} \varphi \, d\varrho(\varphi) = 0$ ist. Dieses Integral ist aber wegen (25) gleich $\mathfrak{M}(g)$. Nach dem erwähnten Bohrschen Satz folgt daraus $g(t) \equiv 0$, was jedoch zu einem Widerspruch führt. Denn der Wertevorrat von $f(t)$ ist, als derjenige einer stetigen Funktion, gewiß zusammenhängend, also, wenn man von den Endpunkten absieht, offenbar mit dem Intervall (17) identisch. Nun liegt α' im Innern dieses Intervalles, also gibt es Werte t_0 mit $f(t_0) = \alpha'$. Wegen (28) folgt daraus $g(t_0) = \alpha'$, also $\alpha' = 0$, während doch $\alpha' > 0$ angenommen wurde.

Damit ist bewiesen, daß $\omega(\varphi)$ auf dem Intervall (17) keine Konstanzstrecke haben kann [(17) kann auch ein Punkt sein, doch ist dann $f(t)$ konstant, also die Behauptung gegenstandslos].

Ist ξ reell, so ist zugleich mit f auch das Produkt $f e_\xi$ fastperiodisch, da $e_\xi = e^{\xi t \sqrt{-1}}$ es ist. Man hat

$$\mathfrak{M}(e_\xi) = 0 \quad \text{für} \quad \xi \gtrless 0 \qquad \text{und} \qquad \mathfrak{M}(e_\xi) = \mathfrak{M}(1) = 1 \quad \text{für} \quad \xi = 0.$$

Denn es ist für $\xi \gtrless 0$

$$(30)\qquad \mathfrak{M}(e_\xi) = \lim_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} e^{\xi t\sqrt{-1}}\,dt = \lim_{T\to+\infty} \frac{\sin \xi T}{\xi T} = 0.$$

Ferner ist

$$(31)\qquad e_\xi e_\eta = e^{\xi+\eta}; \qquad e_0 \equiv 1.$$

Endlich kann der Operator $\mathfrak{M}$ mit gleichmäßigen Grenzübergängen stets vertauscht werden:

aus $\mathfrak{m}(f - f_n) \to 0$ folgt $\mathfrak{M}(f - f_n) \to 0$, $\mathfrak{M}(f_n) \to \mathfrak{M}(f)$,

da $|\mathfrak{M}(f - f_n)| \leqq \mathfrak{m}(f - f_n)$ ist. Ist also $\{\xi_j\}$ irgendeine (also z. B. für $-\infty < \xi < +\infty$ überall dichte) reelle Zahlenfolge, so ist z. B. die (der gleichmäßigen Konvergenz zufolge gewiß fastperiodische) Funktion

$$f(t) = \sum_{j=1}^{\infty} \frac{e_{\xi_j}(t)}{j^2}$$

derart, daß $\mathfrak{M}(f e_{-\xi})$ für $\xi = \xi_j$ und nur für $\xi = \xi_j$ von Null verschieden ausfällt.

Bohr hat nun allgemein bewiesen, daß bei jeder fastperiodischen Funktion $f(t)$ die Funktion $\mathfrak{M}(f e_{-\xi})$ von ξ höchstens für abzählbar unendlich viele (reelle) ξ von Null verschieden sein kann. Es seien $\xi_1, \xi_2, \ldots$ diese singulären Parameterwerte. Die Bezeichnung möge selbstverständlich derart sein, daß aus $j \neq l$ stets $\xi_j \neq \xi_l$ folgt. Die Gesamtheit der singulären Parameter nennen wir *eigentliches* Frequenzsystem von f. Bohr hat — in Verallgemeinerung der Parsevalschen Relation stetiger periodischer Funktionen — die für jede fastperiodische Funktion f gültige „Vollständigkeitsrelation"

$$(32)\qquad \mathfrak{M}(|f|^2) = \sum_j |\mathfrak{M}(f e_{-\xi_j})|^2$$

bewiesen, wobei man über alle eigentlichen Frequenzen von ξ zu summieren hat. Da mit f auch die konjugierte Funktion $\bar{f}$, also auch $|f|^2 = f\bar{f}$ fastperiodisch ist, so folgt daraus einerseits die Konvergenz der unter (32) stehenden Reihe, und andererseits der Eindeutigkeitssatz, daß aus $\mathfrak{M}(f_1 e_{-\xi}) \equiv \mathfrak{M}(f_2 e_{-\xi})$ stets $f_1(t) \equiv f_2(t)$, d. h. aus $\mathfrak{M}(f e_{-\xi}) \equiv 0$ stets $f(t) \equiv 0$ folgt, wenn die Funktionen fastperiodisch sind. Ist nämlich $\mathfrak{M}(f e_{-\xi}) \equiv 0$, so ist wegen (32) $\mathfrak{M}(|f|^2) = 0$, also nach S. 257 auch $|f(t)|^2 \equiv 0$.

Die eigentlichen Frequenzen können im allgemeinen nicht der Größe nach numeriert werden, da sie sogar überall dicht liegen können. —

Unter einem (nicht notwendig eigentlichen) Frequenzsystem $\{\xi_j\}$ von f wollen wir irgendeine Folge von voneinander verschiedenen reellen Zahlen verstehen, die alle eigentlichen Frequenzen von f und darüber hinaus evtl. auch weitere Zahlen enthält. Offenbar ist (32) richtig, auch wenn die ξ_j ein beliebiges Frequenzsystem von f durchlaufen; es ist ja $\mathfrak{M}(f e_{-\xi_j}) = 0$, wenn ξ_j eine uneigentliche Frequenz bedeutet. Anstatt zu sagen, „$\{\xi_j\}$ ist ein Frequenzsystem der fastperiodischen Funktion f", schreibt man, im Anschluß an die Theorie der gewöhnlichen Fourierreihen, die „Äquivalenz"

$$f(t) \sim \sum_j \mathfrak{M}(f e_{-\xi_j}) e_{\xi_j}(t) \tag{33}$$

hin. Ist die Reihe rechterhand (in einer bestimmten, etwa durch die Numerierung der Frequenzen festgelegten Anordnung) gleichmäßig konvergent, so stellt sie die Funktion $f(t)$ dar. Doch kann die Reihe auch divergieren.

Ist

$$g(t) \sim \sum_j \mathfrak{M}(g e_{-\xi_j}) e_{\xi_j}(t), \tag{34}$$

so gilt

$$c_1 f(t) + c_2 g(t) \sim \sum_j [c_1 \mathfrak{M}(f e_{-\xi_j}) + c_2 \mathfrak{M}(g e_{-\xi_j})] e_{\xi_j}(t). \tag{35}$$

Bei der Multiplikation von Äquivalenzen ist eine gewisse Vorsicht geboten, die für die Matrizenzuordnung (siehe weiter unten) sehr wesentlich ist. Wir wollen zunächst eine abkürzende Bezeichnungsweise einführen.

Wir sagen, die Zahlenfolge $\{\zeta_j\}$ bilde einen Modul, wenn mit ζ_j und ζ_l auch $n\zeta_j + m\zeta_l$ der Zahlenfolge angehört, wobei n und $m = 0, \pm 1, \pm 2, \ldots$ sein können. Jede fastperiodische Funktion hat (unendlich viele) Frequenzsysteme $\{\xi_j\}$, die je einen Modul bilden. Man erkennt nämlich nach einem geläufigen, von Cantor herrührenden Diagonalverfahren, daß einer jeden Zahlenfolge $\{\theta_j\}$ ein *abzählbares* Zahlensystem $\{\zeta_j\}$ zugeordnet werden kann derart, daß einerseits jedes θ_j mit einem ζ_j identisch ist, und daß andererseits die ζ_j einen Modul bilden. Kurz, man kann jede Zahlenfolge zu einem Modul ergänzen, und zwar auf unendlich viele Weisen. — Wir sagen, die reelle Zahlenfolge $\{\xi_j\}$ bilde ein *zulässiges* Frequenzsystem von f, wenn die folgenden drei Bedingungen erfüllt sind: 1) $\{\xi_j\}$ ist ein Frequenzsystem von f, so daß also einerseits die beiden Ungleichheiten

$$\xi_j \neq \xi_l, \qquad j \neq l \tag{36}$$

gleichbedeutend sind, und andererseits $\{\xi_j\}$ alle eigentlichen Frequenzen

von f (und evtl. auch weitere Zahlen) enthält; 2) $\{\xi_j\}$ ist ein Modul; 3) die Numerierung ist derart, daß

$$(37) \qquad -\xi_j = \xi_{-j}; \qquad j = 0, \pm 1, \pm 2, \ldots$$

gilt. — Zu 3) ist zu bemerken, daß, wenn $\{\xi_j\}$ ein Modul ist, auch $-\xi_j$ dem Modul angehört, und daß die Null von jedem Modul enthalten wird; wegen (*37*) ist $\xi_0 = 0$. Offenbar hat jede fastperiodische Funktion zulässige Frequenzsysteme. Das eigentliche Frequenzsystem, das — bis auf die Numerierung der Elemente — eindeutig bestimmt ist, braucht kein zulässiges Frequenzsystem zu sein (da es z. B. auch endlich sein kann). Wir sagen, (*33*) sei eine zulässige Darstellung von f, wenn $\{\xi_j\}$ ein zulässiges Frequenzsystem von f ist. Sind

$$(38) \qquad f(t) \sim \sum_i \mathfrak{M}(f e_{-\lambda_i}) e_{\lambda_i}(t), \qquad g(t) \sim \sum_k \mathfrak{M}(f e_{-\mu_k}) e_{\mu_k}(t)$$

zwei beliebige fastperiodische Funktionen, so kann man — nach Hinzufügung von zu verschwindenden Amplituden gehörigen, „parasitären", jedoch für die allgemeine Gültigkeit von (*39*), (*40*) wesentlichen neuen Frequenzen — erreichen, daß sowohl (*36*) als auch (*37*) gilt, wobei $\{\xi_j\}$ ein für beide Funktionen zulässiges Frequenzsystem ist. Wir sagen, (*33*), (*34*) sei eine zulässige Darstellung des Funktionenpaares f, g. Allgemeiner sprechen wir von (stets erreichbaren) zulässigen Darstellungen endlicher oder unendlicher Folgen $f_1, f_2, \ldots$.

Den Bohrschen Multiplikationssatz der Äquivalenzen können wir nun so aussprechen: Ist (*33*), (*34*) eine *zulässige* Darstellung des Funktionenpaares f, g, so darf man diese beiden Äquivalenzen nach den formalen Rechenregeln multiplizieren: es gilt [vgl. (*32*)]

$$(39) \qquad f(t)\, g(t) \sim \sum_j \Big[\sum_l \mathfrak{M}(f e_{\xi_l - \xi_j})\, \mathfrak{M}(g e_{-\xi_l})\Big] e_{\xi_j}(t),$$

mithin

$$(40) \qquad \mathfrak{M}(f g e_{-\xi_j}) = \sum_l \mathfrak{M}(f e_{\xi_l - \xi_j})\, \mathfrak{M}(g e_{-\xi_l}),$$

was, wegen (*2*), nur scheinbar mehr besagt als die aus (*32*) folgende Relation

$$(41) \qquad \mathfrak{M}(fg) = \mathfrak{M}(gf) = \sum_l \mathfrak{M}(f e_{\xi_l})\, \mathfrak{M}(g e_{-\xi_l}).$$

Der Summationszeiger läuft dabei überall von $-\infty$ bis $+\infty$. Die Summen $\sum_l$ sind mit Rücksicht auf (*32*) der Schwarzschen Ungleichheit zufolge absolut konvergent.

Zusammenfassend können wir folgendes behaupten. Es sei $\Phi(z_1, z_2, \ldots, z_n)$ ein Polynom und es sei $f_1(t), \ldots, f_n(t)$ ein System von zulässigen Darstellungen gewisser fastperiodischen Funktionen; dann erhält man eine

zulässige Darstellung der fastperiodischen Funktion $\Phi(f_1, f_2, \ldots, f_n)$, wenn man diesen Ausdruck nach den gewöhnlichen Rechenregeln umordnet. — Wir wollen jetzt insbesondere den Fall $n = 1$ etwas näher betrachten.

Es sei $\{\xi_j^0\}$ das *eigentliche* Frequenzsystem von f, und $\{\eta_j^0\}$ dasjenige von g. Kann jedes η_j^0 aus endlich vielen ξ_j^0 homogen-linear mit ganzen rationalen Koeffizienten zusammengesetzt werden, so wollen wir hierfür kurz $f \succ g$ oder $g \prec f$ schreiben; lies: das eigentliche Frequenzsystem von g ist ein „Teiler" des eigentlichen Frequenzsystems von f. Der vorher erwähnte Satz, wonach die „formale" Ausführung einer ganzen rationalen Operation $\Phi(z)$ auf eine zulässige Darstellung von $f(t)$ eine zulässige Darstellung von $\Phi(f(t))$ liefert, besagt nicht mehr als die Beziehung $f(t) \succ \Phi(f(t))$, in welche keine parasitären Frequenzen eingehen; es ist kaum nötig zu bemerken, daß die Beziehung $\Phi(f(t)) \succ f(t)$ im allgemeinen unrichtig ist. — Hat die gegen f gleichmäßig konvergierende fastperiodische Folge $\{f_n\}$ eine „Frequenzmajorante" F, so daß $F \succ f_n$ für alle n gilt, so gilt auch $F \succ f$. Es sei nämlich ξ irgendeine reelle Zahl. Wegen $\mathfrak{m}(f - f_n) \to 0$ gilt auch $\mathfrak{m}(fe_{-\xi} - f_n e_{-\xi}) \to 0$, da ja $\mathfrak{m}((f - f_n)e_{-\xi}) \leqq \mathfrak{m}(f - f_n)$ ist. Aus $\mathfrak{m}(fe_{-\xi} - f_n e_{-\xi}) \to 0$ folgt aber (vgl. S. 259) die Grenzgleichung $\mathfrak{M}(f_n e_{-\xi}) \to \mathfrak{M}(fe_{-\xi})$. Bezeichnet also ξ eine Zahl, die keine eigentliche Frequenz von F ist, so daß wegen $F \succ f_n$ auch $\mathfrak{M}(f_n e_{-\xi}) = 0$ gilt, so muß auch $\mathfrak{M}(fe_{-\xi})$ verschwinden, d. h., ξ ist keine eigentliche Frequenz von f, w. z. b. w.

Bewegt sich t von $-\infty$ bis $+\infty$, so beschreibt $z = f(t)$ eine Kurve in der Ebene der komplexen Variable $z = x + \sqrt{-1}\,y$. Die Gesamtheit der Punkte, welchen die Kurve beliebig nahe kommen kann (ohne sie notwendig zu erreichen), möge abgeschlossener Wertevorrat von f genannt und mit $\mathfrak{w}(f)$ bezeichnet werden. Ist $f(t)$ reellwertig, so ist $\mathfrak{w}(f)$ mit dem Intervall (*17*) identisch. Ein anderes Beispiel ist dasjenige einer Lissajousschen Kurve $x = A_1 \cos \varrho_1 t$, $y = A_2 \cos \varrho_2 t$, durch welche ein Rechtecksbereich „ausgefüllt" wird, so daß $\mathfrak{w}(f)$ dieser Rechtecksbereich ist, trotzdem die Kurve keine Peanokurve ist, indem sie nicht durch *jeden* Punkt des Rechtecks mindestens einmal hindurchgeht.

Es sei $\Psi(z)$ eine evtl. nur auf der Punktmenge $\mathfrak{w}(f)$ erklärte, auf $\mathfrak{w}(f)$ stetige Funktion. Da $\mathfrak{w}(f)$ beschränkt und abgeschlossen ist, so gibt es nach einer bekannten Verallgemeinerung des Weierstraßschen polynomischen Approximationssatzes eine Folge $\Phi_1(z)$, $\Phi_2(z), \ldots$ von ganzen rationalen Funktionen derart, daß auf $\mathfrak{w}(f)$ gleichmäßig die Grenzgleichung $\Phi_n(z) \to \Psi(z)$ gilt. Offenbar gilt

also $\Phi_n(f(t)) \to \Psi(f(t))$ gleichmäßig für alle t. Mithin ist $\Psi(f(t))$ fastperiodisch, (S. 252), und zwar ist nach dem vorangehenden $f(t) \succ \Psi((t))$. Daher ist jedes zulässige Frequenzsystem von $f(t)$ auch für $\Psi(f(t)$ zulässig. Um also eine (durch die Multiplikationsregel gebotene) zulässige Darstellung des Funktionen*paares* $f(t)$, $\Psi(f(t))$ zu erhalten, brauchen wir zu einem zulässigen Frequenzsystem von $f(t)$ keine weiteren parasitären Frequenzen hinzuzufügen. Diese Bemerkung wird für unsere Matrizenzuordnung von einigem Interesse sein. Es wird sich dabei vor allem — nämlich bei der Behandlung der Reziproke oder Resolvente und damit des Spektrums — um die gebrochene rationale Funktion $\Psi(z) = z^{-1}$ handeln.

Die untere Grenze von $|f(t)|$ haben wir mit $\mathfrak{n}(f)\ (\geqq 0)$ bezeichnet, so daß $\mathfrak{n}(f) > 0$ dann und nur dann gilt, wenn $f(t)$ der Null nicht beliebig nahe kommen kann. Ist also $\mathfrak{n}(f) > 0$, so ist die Funktion z^{-1} auf der Punktmenge $\mathfrak{w}(f)$ stetig, mithin ist dann $(f(t))^{-1}$ fastperiodisch und es gilt $f(t) \succ (f(t))^{-1}$. Ist aber $\mathfrak{n}(f) = 0$, so ist die Funktion $(f(t))^{-1}$, wobei $f \not\equiv 0$ vorausgesetzt werden kann, nicht beschränkt, um so weniger fastperiodisch. Ist also $(f(t))^{-1}$ überhaupt fastperiodisch, so ist es auch $\prec f(t)$.

Wir werden endlich die fast selbstverständliche Bemerkung benötigen, daß stets $f(t) \succ \bar{f}(t)$, also auch $\bar{f}(t) \succ f(t)$ gilt, wobei $\bar{f}(t) = \overline{f(t)}$ die konjugierte Funktion bedeutet.

Es sei nun $f(t)$ irgendeine fastperiodische Funktion und

$$f(t) \sim \sum_{j=-\infty}^{+\infty} \mathfrak{M}(f e_{-\xi_j})\, e_{\xi_j}(t) \tag{42}$$

eine der zulässigen Darstellungen von f. Wir ordnen dieser Darstellung von f die Matrix

$$\mathfrak{B}(f) = \| b_{ik}(f) \| = \| \mathfrak{M}(f e_{\xi_i - \xi_k}) \|; \qquad i, k = 0, \pm 1, \ldots \tag{43}$$

als die (zu $\{\xi_j\}$ gehörige) „fastperiodische Matrix von f“ zu. Die Zeiger laufen ganzzahlig nicht von 1, sondern von $-\infty$ bis $+\infty$, so daß nicht eine Viertelebene, sondern die ganze Ebene besetzt ist:

$$\mathfrak{B}(f) = \left\| \begin{matrix} \ddots & & \ddots & & \\ & \mathfrak{M}(f e_{\xi_{-1}-\xi_{-1}}) & \mathfrak{M}(f e_{\xi_{-1}-\xi_0}) & \mathfrak{M}(f e_{\xi_{-1}-\xi_1}) & \\ \ddots & \mathfrak{M}(f e_{\xi_0-\xi_{-1}}) & \mathfrak{M}(f e_{\xi_0-\xi_0}) & \mathfrak{M}(f e_{\xi_0-\xi_1}) & \ddots \\ & \mathfrak{M}(f e_{\xi_1-\xi_{-1}}) & \mathfrak{M}(f e_{\xi_1-\xi_0}) & \mathfrak{M}(f e_{\xi_1-\xi_1}) & \\ & & \ddots & & \ddots \end{matrix} \right\|;$$

doch gilt die allgemeine Theorie auch bei diesen Matrizen; man braucht nur bei den Abschnitten nicht von 1 bis n, sondern von $-n$ bis n zu summieren, die unendliche Mannigfaltigkeit E durch $\sum_{i=-\infty}^{+\infty} |x_i|^2 = 1$, ihren Abschnitt $\mathsf{E}_{[n]}$ durch $\sum_{i=-n}^{+n} |x_i|^2 = 1$ festzulegen usf. In der Hauptdiagonale $i = k$ steht wegen $e_{\xi_i - \xi_i} = e_0 = 1$ stets der Mittelwert von f,

$$(44) \qquad b_{ii}(f) = \mathfrak{M}(f).$$

Da wegen (*37*) $\xi_0 = 0$ ist, so ist die nullte Zeile mit der Folge der unter (*42*) angeschriebenen Amplituden von f identisch. Ist also $\mathfrak{B}(f)$ die Nullmatrix, so ist wegen (*32*) auch $\mathfrak{M}(|f|^2) = 0$, und daher $f(t) \equiv 0$. Ist $f = c =$ Konstante, so ist $\mathfrak{B}(f)$ wegen (*30*) die Diagonalmatrix $c\,\mathfrak{E}$, wobei $\mathfrak{E} = \mathfrak{B}(1)$ ist, so daß die Zeiger auch in $\mathfrak{E}$ von $-\infty$ bis $+\infty$ laufen. Ist umgekehrt $\mathfrak{B}(f)$ eine Diagonalmatrix, etwa $= \| \Gamma_{ii}\,\delta_{ik} \|$, so muß zunächst $\Gamma_{ii} = \Gamma_{00}$, also $\mathfrak{B}(f) = \Gamma_{00}\,\mathfrak{E}$ gelten. Da aber auch $\mathfrak{B}(\Gamma_{00}) = \Gamma_{00}\,\mathfrak{E}$ ist, so folgt daraus $\mathfrak{B}(f - \Gamma_{00}) = \mathfrak{O}$ und daher, wie man sich leicht überzeugt, $f(t) - \Gamma_{00} \equiv 0$, so daß $f(t)$ konstant ist.

Ist die Folge $\{\xi_j\}$ äquidistant, $\xi_j = j$, so gehen die $\mathfrak{B}(f)$ in die zu eigentlichen (periodischen) Fourierreihen gehörigen Laurentmatrizen von Toeplitz über. In diesem Falle ist $\xi_i - \xi_k = i - k$, so daß nicht nur die Hauptdiagonale, sondern auch jede „Nebendiagonale" (= zu der Hauptdiagonale parallele Gerade) aus dem auf der nullten Zeile gelegenen Element der Nebendiagonale durch Verschiebung längs der Nebendiagonale hervorgeht. Aber auch im allgemeinen, nicht äquidistanten Fall ist jedes Element der Matrix eine der unter (*33*) angeschriebenen Amplituden oder — und meistens — gleich Null. In diesem Sinne können also keine „Resonanzglieder" auftreten. Die Numerierung der ξ_j ist, sobald die Bedingung (*37*) erfüllt ist, des näheren willkürlich (wir haben uns die Numerierung irgendwie fest zu wählen). Der Leser möge sich die Struktur der fastperiodischen Matrizen bei passender Numerierung in dem einfachsten nicht äquidistanten Falle klarmachen, wenn nämlich $f(t)$ „bedingtperiodisch" ist, d. h. wenn das eigentliche Frequenzsystem von f in einem Modul liegt, die durch endlich viele (> 1) linear-unabhängige Zahlen bestimmt wird (z. B. bei der auf S. 262 erwähnten Lissajousschen Kurve).

Aus (*43*) folgt leicht, daß $\mathfrak{B}(f)$ dann und nur dann von Hermiteschem Typus ist, wenn $f(t)$ reellwertig ist. Ferner ist $\mathfrak{B}(f)$ dann und nur dann symmetrisch, $b_{ik}(f) = b_{ki}(f)$, wenn $f(t)$ gerade, $\equiv f(-t)$ ist [und ebenso stellt $f(t) \equiv -f(-t)$ die notwendige und

hinreichende Bedingung der schiefen Symmetrie $b_{ik}(f) = -b_{ki}(f)$ dar]. Mithin kann (33) dann und nur dann in der Gestalt einer Kosinusreihe

$$\sum_{j=1}^{\infty} A_j \cos \xi_j t$$

mit reellen Amplituden A_j geschreiben werden, wenn $\mathfrak{B}(f)$ reell und symmetrisch ist. — Allgemein ist wegen $\bar{e}_\xi(t) = e_{-\xi}(t)$ die begleitende Matrix $\mathfrak{B}^*(f)$ von $\mathfrak{B}(f)$ offenbar

$$\| \overline{\mathfrak{M}(f e_{\xi_i - \xi_k})} \|' = \| \mathfrak{M}(\bar{f} e_{\xi_k - \xi_i}) \|' = \| \mathfrak{M}(\bar{f} e_{\xi_i - \xi_k}) \|,$$

so daß

$$\mathfrak{B}^*(f) = \mathfrak{B}(\bar{f}) \tag{45}$$

ausfällt.

Wir wollen jetzt beweisen, daß $\mathfrak{B}(f)$ beschränkt ist. Es bezeichne zu diesem Ende

$$\overset{0}{\mathfrak{x}}_{(n)} = (\overset{0}{x}_{-n}, \overset{0}{x}_{-n+1}, \ldots, \overset{0}{x}_{-1}, \overset{0}{x}_0, \overset{0}{x}_1, \ldots, \overset{0}{x}_n) \tag{46}$$

irgendeinen Punkt auf $\mathsf{E}_{[n]}$, und $h(t)$ die endliche fastperiodische Reihe

$$h(t) = \sum_{j=-\infty}^{+\infty} x_j^0 e_{\xi_j}(t); \qquad x_j^0 = 0 \quad \text{für} \quad |j| > n, \tag{47}$$

so daß, wegen (32), und da $\overset{0}{\mathfrak{x}}_{(n)}$ auf $\mathsf{E}_{[n]}$ liegt,

$$\mathfrak{M}(|h|^2) = 1 \tag{48}$$

gilt. Wegen $x_j^0 = \mathfrak{M}(h e_{-\xi_j})$ folgt aus (40) offenbar

$$\sum_{i=-\infty}^{+\infty} \mathfrak{M}(f e_{\xi_i - \xi_k}) x_i^0 = \sum_{i=-\infty}^{+\infty} \mathfrak{M}(f e_{\xi_i - \xi_k}) \mathfrak{M}(h e_{-\xi_i}) = \mathfrak{M}(f h e_{-\xi_k}), \tag{49}$$

also ist

$$\left| \sum_{k=-\infty}^{+\infty} \sum_{i=-\infty}^{+\infty} \mathfrak{M}(f e_{\xi_i - \xi_k}) \bar{x}_k^0 x_i^0 \right|^2$$

$$\leqq \left(\sum_{k=-\infty}^{+\infty} |\bar{x}_k^0|^2 \right) \sum_{k=-\infty}^{+\infty} \left| \sum_{i=-\infty}^{+\infty} \mathfrak{M}(f e_{\xi_i - \xi_k}) x_i^0 \right|^2 = \sum_{k=-\infty}^{+\infty} |\mathfrak{M}(f h e_{-\xi_k})|^2$$

$$= \mathfrak{M}(|fh|^2) \leqq \mathfrak{M}([\mathfrak{m}(f)]^2 |h|^2) = [\mathfrak{m}(f)]^2 \mathfrak{M}(|h|^2) = [\mathfrak{m}(f)]^2,$$

wegen (40) und (48). Die Kopplungsform von $\mathfrak{B}(f)$ auf $\mathsf{E}_{[n]}$ ist daher absolut $\leqq \mathfrak{m}(f)$. Mithin ist $\mathfrak{B}(f)$ beschränkt und man hat

$$\mathsf{M}(\mathfrak{B}(f)) \leqq \mathfrak{m}(f). \tag{50}$$

Es läßt sich sogar mehr als dies behaupten, doch können wir darauf nicht eingehen.

Es sei $g(t)$ irgendeine fastperiodische Funktion, wobei, wie wir wissen, angenommen werden kann, daß sie zugleich mit (33) auf eine zulässige Weise dargestellt ist,

$$(51) \qquad g(t) \sim \sum_{j=-\infty}^{+\infty} \mathfrak{M}(g e_{-\xi_j}) e_{\xi_j}(t).$$

Aus (35) folgt

$$(52) \qquad \mathfrak{B}(f) + \mathfrak{B}(g) = \mathfrak{B}(f+g).$$

Das Matrizenprodukt $\mathfrak{B}(f)\mathfrak{B}(g)$ kann unter Benutzung von (40) unmittelbar berechnet werden. Es ergibt sich offenbar

$$(53) \qquad \mathfrak{B}(f)\mathfrak{B}(g) = \mathfrak{B}(fg).$$

Endlich ist, wie wir wissen,

$$(54) \qquad \mathfrak{B}(1) = \mathfrak{E}.$$

Mithin gilt

$$(55) \qquad F(\mathfrak{B}(f)) = \mathfrak{B}(F(f)),$$

wobei F irgendeine ganze rationale Funktion bedeutet. Ist F eine willkürliche, auf $\mathfrak{w}(f)$ erklärte stetige Funktion, so ist $F(f(t))$ eindeutig erklärt und fastperiodisch. Wir wollen in diesem Falle (55) als die allgemeine Definition der Matrix $F(\mathfrak{B}(f))$ betrachten. — Ist $g_1, g_2, \ldots$ eine Folge von fastperiodischen Funktionen, die auf eine auf für g zulässige Weise dargestellt sind und gleichmäßig gegen g konvergieren, so konvergiert die Kopplungsform $\Phi(\mathfrak{B}(g_n))$ von $\mathfrak{B}(g_n)$ auf E gleichmäßig gegen $\Phi(\mathfrak{B}(g))$, da wegen (50) auf E

$$|\Phi(\mathfrak{B}(g)) - \Phi(\mathfrak{B}(g_n))| = |\Phi(\mathfrak{B}(g-g_n))| \leqq \mathsf{M}(\mathfrak{B}(g-g_n)) \leqq \mathfrak{m}(g-g_n) \rightarrow 0$$

gilt. — Unsere Zuordnung besitzt demnach alle Eigenschaften der Klassen von funktionalen Gruppenmatrizen. Dies beruht sehr wesentlich auf dem Umstand, daß nur zulässige Darstellungen betrachtet werden. Sonst könnte nämlich (40), also auch (53) ungültig sein. Die parasitären ξ_j, für welche $\mathfrak{M}(f e_{-\xi_j}) = 0$ ist, benötigt man hierbei gewissermaßen zu einem „Platzmachen" (Trennung).

Mit Rücksicht auf (53) sind $\mathfrak{B}(f)$ und $\mathfrak{B}(g)$ offenbar vertauschbare Matrizen. Andererseits ist nach (45) die begleitende Matrix von $\mathfrak{B}(f)$ gleich $\mathfrak{B}(\bar{f})$, also eine Matrix $\mathfrak{B}(g)$. Mithin ist die Matrix $\mathfrak{B}(f)$ normal. Ihre Norm berechnet sich aus (45), (53) zu $\mathfrak{B}(|f|^2)$. Kann f der Null nicht beliebig nahe kommen, so daß $\mathfrak{n}(f) > 0$ gilt, so ist f^{-1} fastperiodisch und $\prec f$, so daß wir in (53) $g = f^{-1}$ setzen dürfen. Mit

Rücksicht auf (*54*) folgt sodann

$$(56)\qquad \mathfrak{B}(f)\,\mathfrak{B}\left(\frac{1}{f}\right) = \mathfrak{E}.$$

Da f normal ist, so ist der pathologische Fall 3 (S. 138) ausgeschlossen. Ist also $\mathfrak{u}(f) > 0$, so ist $\mathfrak{B}\left(\frac{1}{f}\right)$ eine sowohl hintere als auch vordere, beschränkte Reziproke von $\mathfrak{B}(f)$.

Ist $\mathfrak{u}(f) = 0$, so können wir zunächst nur behaupten, daß die Matrix

$$(57)\qquad \mathfrak{B}\left(\frac{1}{f}\right) = \left\| \lim_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} \frac{e_{\xi_i-\xi_k}(t)}{f(t)} \right\|$$

entweder überhaupt nicht gebildet werden kann, oder, wenn sie gebildet werden könnte (indem sich $f(t)$ nur „langsam" der Null nähert), gewiß keine fastperiodische Matrix darstellt, da ja f^{-1} wegen $\mathfrak{u}(f) = 0$ nicht fastperiodisch ist. Wir können also zunächst nicht behaupten, daß $\mathfrak{B}(f)$ *überhaupt keine* beschränkte Reziproke hat. Dies ist vielmehr ein tiefliegender Satz, der die spektrale Untersuchung von $f(t)$ im wesentlichen zu beherrschen vermag, und den wir nur auf einem scheinbaren Umwege beweisen können. Wir bestimmen nämlich eigentlich *alle* matrizentheoretisch singulären Stellen von $\mathfrak{B}(f)$ für den Fall, wo $\mathfrak{B}(f)$ Hermitesch ist und führen sodann die Entscheidung der Frage, ob in dem allgemeinen Falle, d. h. wenn $\mathfrak{B}(f)$ nicht mehr Hermitesch ist, $\mathfrak{B}(f)$ eine beschränkte Reziproke hat oder nicht, mit Hilfe der Toeplitzschen Kriterien auf den Hermiteschen Fall zurück.

Es sei zunächst die fastperiodische Funktion $f(t)$ beliebig und es bezeichne λ irgendeine, außerhalb der Punktmenge $\mathfrak{w}(f)$ gelegene Konstante. Dann kann $\lambda - f(t)$ der Null nicht beliebig nahe kommen, so daß $(\lambda - f(t))^{-1}$ fastperiodisch ist. Ferner gilt $f(t) \supset \lambda - f(t) \supset (\lambda - f(t))^{-1}$, da die Konstante λ zur Frequenz Null gehört, die in jedem zulässigen Frequenzsystem enthalten ist. Offenbar ist also $\mathfrak{B}((\lambda - f)^{-1})$ eine, also die einzige beschränkte Reziproke von

$$(58)\qquad \mathfrak{B}(\lambda - f) = \mathfrak{B}(\lambda) - \mathfrak{B}(f) = \lambda\,\mathfrak{E} - \mathfrak{B}(f),$$

d. h. die Resolvente $\|R_{ik}(\lambda)\|$ von $\mathfrak{B}(f)$, so daß alle außerhalb $\mathfrak{w}(f)$ gelegenen λ matrizentheoretisch gewiß regulär sind. Es gilt

$$(59)\qquad R_{ik}(\lambda) = \mathfrak{M}\left(\frac{e_{\xi_i-\xi_k}}{\lambda - f}\right),$$

also insbesondere $R_{00}(\lambda) =$

$$(60)\qquad \mathfrak{M}\left(\frac{1}{\lambda - f}\right) = \lim_{T\to+\infty} \frac{1}{2T} \int_{-T}^{T} \frac{dt}{\lambda - f(t)}.$$

Es sei nun $f(t)$ reellwertig, d. h. $\mathfrak{B}(f)$ von Hermiteschem Typus, und es bezeichne $\|\sigma_{ik}(\mu)\|$ die Spektralmatrix von $\mathfrak{B}(f)$, so daß

$$(61)\qquad R_{ik}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\sigma_{ik}(\mu)}{\lambda - \mu},$$

also wegen (59) und mit Rücksicht auf die Stieltjessche Umkehrformel

$$(62)\qquad R_{ii}(\lambda) = R_{00}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\sigma_{00}(\mu)}{\lambda - \mu},$$

und

$$(63)\qquad \sigma_{ii}(\mu) = \sigma_{00}(\mu); \qquad i = \pm 1, \pm 2, \ldots$$

gilt. Ist λ zunächst etwa reell, so läßt sich (60) nach einer fast wortgetreuen Wiederholung der bei dem Beweis von (25) angewandten Schlußweise in

$$(64)\qquad R_{00}(\lambda) = \int_{-\infty}^{+\infty} \frac{d\omega(\mu)}{\lambda - \mu}$$

umrechnen, wobei $\omega(\mu)$ dieselbe Bedeutung wie in (25) hat, d. h. die Dichtefunktion von $f(t)$ bezeichnet. Aus Gründen der analytischen Fortsetzung gilt (64) offenbar auch dann, wenn λ nicht reell ist, jedoch stets außerhalb $\mathfrak{w}(f)$ liegt. Aus (62), (64) folgt nach der Stieltjesschen Umkehrformel, daß an allen Stetigkeitsstellen von $\omega(\mu)$

$$(65)\qquad \omega(\mu) = \sigma_{00}(\mu)$$

gilt, sofern man nötigenfalls von einer additiven Konstante absieht. Diese additive Konstante ist aber gleich Null, da $\sigma_{00}(\mu)$ nach der allgemeinen Theorie, $\omega(\mu)$ aber wegen (23) für hinreichend große μ gleich Eins ist. Nun hat $\omega(\mu)$ auf dem Intervall

$$(66)\qquad \mathfrak{n}(f) \leqq \mu \leqq \mathfrak{m}(f)$$

keine Konstanzstrecke (S. 258). Wegen (65) folgt daher aus dem auf S. 98 bewiesenen Satz, daß alle Punkte dieses Intervalles singuläre Stellen der Funktion $R_{00}(\lambda)$, also (nach S. 144) erst recht solche von $\mathfrak{B}(f)$ sind, mithin gewiß dem Spektrum von $\mathfrak{B}(f)$ angehören. Andererseits ist das Intervall (66) nach Definition mit $\mathfrak{w}(f)$ identisch (S. 262), und alle außerhalb $\mathfrak{w}(f)$ liegenden λ sind, wie wir auf S. 267 gesehen haben, reguläre Stellen von $\mathfrak{B}(f)$, liegen also außerhalb des Spektrums.

Folglich ist das Spektrum von $\mathfrak{B}(f)$ mit dem Intervall (*66*) identisch. Nun sind aber die beiden Endpunkte des Spektrums von $\mathfrak{B}(f)$ nach dem auf S. 146 ff. bewiesenen Satz die Punkte $\mu = \mathsf{n}(\mathfrak{B}(f))$ und $\mu = \mathsf{m}(\mathfrak{B}(f))$, wobei $\mathsf{n}(\mathfrak{B}(f))$ die untere, $\mathsf{m}(\mathfrak{B}(f))$ die obere Grenze aller Abschnittseigenwerte von $\mathfrak{B}(f)$ bedeuten, sie sind also die zu $n \to +\infty$ gehörige untere bzw. obere Grenze des Minimums bzw. des Maximums des n-ten Abschnittes der Kopplungsform von $\mathfrak{B}(f)$ auf $\mathsf{E}_{[n]}$, also einfach die untere bzw. obere Grenze der Kopplungsform von $\mathfrak{B}(f)$ auf E. Nun ist der Wertevorrat der Kopplungsform auf E ein (zusammenhängendes) Intervall. Mithin ist der Wertevorrat $\mathfrak{w}(f)$ der Funktion $f(t)$ identisch mit dem abgeschlossenen Wertevorrat $\mathsf{W}(\mathfrak{B}(f))$ der Kopplungsform von $\mathfrak{B}(f)$ auf E. Dieser Satz ist weniger scharf als der vorige, der die Identität von $\mathfrak{w}(f)$ mit dem Spektrum aussagt. Es wäre ja denkbar, daß das Spektrum zwischen seinen beiden Endpunkten

$$(67) \qquad n(f) = \mathsf{n}(\mathfrak{B}(f)), \qquad m(f) = \mathsf{m}(\mathfrak{B}(f))$$

Lücken aufweist (vgl. S. 148). Daß dem nicht so ist, folgt daraus, daß $\mathfrak{w}(f)$ zusammenhängend und mit dem Spektrum identisch ist.

Es war bisher vorausgesetzt, daß $\mathfrak{B}(f)$ Hermitesch ist. Den allgemeinen Fall können wir unter Benutzung eines methodischen Gedankens von der Toeplitzschen Reziprokentheorie auf diesen zurückführen. Es sei also $f(t)$ eine beliebige fastperiodische Funktion. $\mathfrak{B}(f)$ ist normal, also ist der Fall **3** (S. 138) ausgeschlossen; $\mathfrak{B}(f)$ hat nach S. 138 dann und nur dann gar keine beschränkte Reziproke, wenn der Fall **2** vorliegt, d. h. wenn die Kopplungsform der Norm von $\mathfrak{B}(f)$ auf E der Null beliebig nahe kommen kann. Die Norm von $\mathfrak{B}(f)$ ist aber nach S. 266 die Matrix $\mathfrak{B}(f)\,\mathfrak{B}(\bar{f}) = \mathfrak{B}(|f|^2)$. Da sie Hermitesch ist, so ist nach dem Vorhergehenden $\mathsf{W}(\mathfrak{B}(|f|^2)) = \mathfrak{w}(|f|^2)$, so daß $\Phi(\mathfrak{B}(|f|^2))$ auf E der Null dann und nur dann beliebig nahe kommen kann, wenn $|f(t)|^2$, d. h. $f(t)$ es tut. M. a. W., dann und nur dann, wenn $\mathfrak{n}(f) = 0$ ist. Ist also $\mathfrak{n}(f) = 0$, so ist nicht nur $\mathfrak{B}(f^{-1})$ keine beschränkte Reziproke von $\mathfrak{B}(f)$, sondern es gibt gar keine beschränkte Reziproke, w. z. b. w.

Die Matrix (*43*) hat demnach dann und nur dann keine beschränkte Reziproke, wenn die Funktion $\lambda - f(t)$ der Null, d. h. wenn $f(t)$ der Konstanten λ beliebig nahe kommen kann, m. a. W., wenn λ auf $\mathfrak{w}(f)$ liegt. Anders ausgedrückt, die Menge der matrizentheoretisch singulären Stellen von $\mathfrak{B}(f)$ ist mit dem abgeschlossenen Wertevorrat $\mathfrak{w}(f)$ der Funktion $f(t)$ identisch. Im Hermiteschen Spezialfall haben wir den Satz bereits vorher bewiesen.

Wir können zum Schluß den Grenzwertsatz (*10*), (*15*), mithin die Existenz der Dichtefunktion $\omega(\varphi)$, sowie die Beziehung (*64*) auf eine von

den Betrachtungen auf S. 252 ff. unabhängige Weise begründen. Zu diesem Ende genügt es zu bemerken, daß wegen (*9*), analog zu (*18*), die identische Umformung

$$(68)\qquad \frac{1}{2T}\int_{-T}^{T}\frac{dt}{\lambda - f(t)} = \int_{-\infty}^{+\infty}\frac{d\nu_T(\varphi)}{\lambda - \varphi}$$

gilt (etwa für hinreichend große λ). Das übrige ergibt sich hieraus wegen (*60*) einfach aus dem Grommerschen Fundamentalsatz (S. 102 ff.) [die Definition (*9*) ist bei (*68*) für alle φ in Anspruch genommen worden].

Ist $\mathfrak{B}(f)$ unitär, so sind, wie wir wissen (S. 212), alle singulären Stellen der Matrix vom Betrage Eins, also muß dann die Kurve $z = f(t)$, die auch auf einen Punkt zusammenschrumpfen kann, auf dem Rande $|z| = 1$ des Einheitskreises verlaufen. Ist umgekehrt diese Bedingung erfüllt, so daß $|f(t)| \equiv 1$ gilt, so ist $\bar{f} = f^{-1}$, also wegen (*45*) und (*56*) $\mathfrak{B}(f)\,\mathfrak{B}^*(f) = \mathfrak{E}$, d. h. $\mathfrak{B}(f)$ ist dann unitär (da es gewiß normal ist). Ist z. B. $F(t) = \exp\left(\sqrt{-1}\,f(t)\right)$ und $f(t)$ reellwertig und fastperiodisch, so ist die Matrix $\mathfrak{B}(F)$ gewiß unitär. Aber auch die Matrix $\mathfrak{B}(e_1)$ ist unitär [nicht aber die Matrix $\mathfrak{A}$ auf S. 138; der Unterschied rührt davon her, daß es in $\mathfrak{B}(e_1)$ keine erste Kolonne gibt]. — Die Bemerkung, daß $\mathfrak{B}(F)$ unitär ist, bildet übrigens nur einen Spezialfall der folgenden. Ist $\mathfrak{H}$ irgendeine beschränkte Hermitesche Matrix, so ist die Matrix

$$\mathfrak{F} = \exp\left(\sqrt{-1}\,\mathfrak{H}\right) = \sum_{m=0}^{\infty}\frac{\left(\sqrt{-1}\,\mathfrak{H}\right)^m}{m!}$$

beschränkt und gewiß unitär. Die Beschränktheit folgt aus der bei (230) gemachten Bemerkung mit Rücksicht auf (231) unmittelbar. Man hat ferner wegen $\mathfrak{H} = \mathfrak{H}^*$

$$\left[\left(\sqrt{-1}\,\mathfrak{H}\right)^{2n}\right]^* = (-1)^n\left(\mathfrak{H}^{2n}\right)^* = (-1)^n\,\mathfrak{H}^{2n} = \left(-\sqrt{-1}\,\mathfrak{H}\right)^{2n},$$

also wegen $\overline{\sqrt{-1}} = -\sqrt{-1}$

$$\begin{aligned}\left[\left(\sqrt{-1}\,\mathfrak{H}\right)^{2n+1}\right]^* &= \left(\sqrt{-1}\,\mathfrak{H}\right)^*\left[\left(\sqrt{-1}\,\mathfrak{H}\right)^{2n}\right]^* = \left(\overline{\sqrt{-1}}\,\mathfrak{H}\right)\left[\left(\sqrt{-1}\,\mathfrak{H}\right)^{2n}\right]\\ &= \left(-\sqrt{-1}\,\mathfrak{H}\right)^{2n+1}\end{aligned}$$

und daher offenbar $\mathfrak{F}^* = \exp\left(-\sqrt{-1}\,\mathfrak{H}\right)$. Da die Funktionalgleichung $(\exp\mathfrak{A}_1)(\exp\mathfrak{A}_2) = \exp(\mathfrak{A}_1 + \mathfrak{A}_2)$ für beschränkte und vertauschbare $\mathfrak{A}$ aus der Potenzreihe heraus unmittelbar verifiziert werden kann, und da andererseits $\exp\mathfrak{O} = \mathfrak{E}$ gilt, so ist daher $\mathfrak{F}\,\mathfrak{F}^* = \mathfrak{F}^*\,\mathfrak{F} = \mathfrak{E}$, w. z. b. w.

Wir haben σ_{00} auf S. 268 dadurch bestimmt, daß wir in das unter (*60*) stehende „Integral" $\mathfrak{M}$ einfach f als unabhängige Variable einführten. Durch sinngemäße Verallgemeinerung[1]) der vorher durchgeführten Betrachtungen ließe sich f etwa auch in $\mathfrak{M}(gf)$ als unabhängige Variable einführen, wobei g auch von t abhängt, und es ließe sich dann ebenso σ_{ik} auch für $i \neq k$ berechnen; doch würde dies hier zu weit führen.

Da der Operator $\mathfrak{M}$ mit gleichmäßigen Grenzübergängen vertauschbar ist, so kann (*60*) wegen $|\mathfrak{M}(f^n)| \leqq [\mathfrak{m}(f)]^n$ für hinreichend große $|\lambda|$ nach Potenzen von λ einfach dadurch entwickelt werden, daß man die gewöhnlichen Rechenregeln anwendet (vgl. S. 101). Da dasselbe auch für (*64*) gilt, so ergibt der Vergleich, daß

$$\mathfrak{M}(f^n) = \int_{-\infty}^{+\infty} \mu^n \, d\omega(\mu); \qquad n = 0, 1, 2, \ldots,$$

also

$$\mathfrak{M}(G(f)) = \int_{-\infty}^{+\infty} G(\mu) \, d\omega(\mu)$$

gilt, wobei $G(z)$ zunächst ein Polynom ist, von welchem man jedoch offenbar zu willkürlichen Funktionen übergehen kann, die auf $\mathfrak{w}(f)$ erklärt und stetig sind.

Wir haben auf S. 181 eine Regel kennengelernt, welche die Spektralmatrix von $\mathfrak{A}^{-1}$ aus derjenigen von $\mathfrak{A}$ zu berechnen gestattet, nämlich die Formel (310); diese, auf $\sigma_{00} = \omega$ angewandt, ergibt offenbar die Dichtefunktion von f^{-1} (sofern $\mathfrak{A}^{-1}$ überhaupt vorhanden); von den Unstetigkeitsstellen ist dabei abzusehen. — Der einfachste Fall ist der, wo $f(t)$ auf $\mathfrak{w}(f)$ „gleichverteilt", nämlich $\omega(\mu)$ linear ist:

$$\omega(\mu) = \frac{\mu - \boldsymbol{n}(f)}{\boldsymbol{m}(f) - \boldsymbol{n}(f)} \quad \text{für} \quad \boldsymbol{n}(f) \leqq \mu \leqq \boldsymbol{m}(f),$$

$$\omega(\mu) = 0 \quad \text{für} \quad \mu < \boldsymbol{n}(f), \qquad \omega(\mu) = 1 \quad \text{für} \quad \mu > \boldsymbol{m}(f).$$

Die Regel (310) kann, wie auf S. 181 erwähnt wurde, auch auf andere etwa stückweise monotone Funktionen von $\mathfrak{A}$ [d. h. von f] übertragen werden. Grundsätzlich sehr einfach ist auch der Fall, wo es sich um eine willkürliche fastperiodische Funktion von $\mathfrak{A}$ [d. h. von f] handelt.

Und zum Schluß noch eine Bemerkung. Wir hatten keine Regel angegeben, welche $\sigma_{ik}(\mu)$ auch für $i \neq k$ zu berechnen gestatten würde, wir haben vielmehr nur $\sigma_{00}(\mu)$ bestimmt und aus den Nichtkonstanzstellen von σ_{00}, wie es sich nachträglich ergeben hat, dennoch das

[1]) Für eigentliche Integrale vgl. F. Riesz, Les systèmes d'équations linéaires, Paris (1913), S. 153.

ganze Spektrum ablesen können, während das Spektrum, kurz gesprochen, aus den Nichtkonstanzstellen von *allen* σ_{ik} besteht (vgl. nämlich S. 178). Woher kommt es nun, daß jetzt sich das ganze Spektrum aus einer einzigen Funktion ablesen ließ?

Es sei $\|\sigma_{ik}(\mu)\|$ die Spektralmatrix irgendeiner beschränkten Hermiteschen Matrix $\mathfrak{A}$. Die Kopplungsform $\Phi(\|\sigma_{ik}(\mu)\|)$ ist, wie wir wissen, in jedem Punkte von E eine nicht abnehmende Funktion von μ, d. h. die Hermitesche Matrix $\|\underset{\mathrm{a}}{\Delta}\sigma_{ik}\|$ ist bei jedem a nichtnegativ definit, also sind es auch ihre Abschnitte, so daß wegen (60)

$$|\underset{\mathrm{a}}{\Delta}\sigma_{ik}|^2 \leqq \underset{\mathrm{a}}{\Delta}\sigma_{ii}\cdot\underset{\mathrm{a}}{\Delta}\sigma_{kk} \qquad (\underset{\mathrm{a}}{\Delta}\sigma_{ii} \geqq 0,\ \underset{\mathrm{a}}{\Delta}\sigma_{kk} \geqq 0)$$

für alle i und k gilt. Ist also auf a jedes Diagonalelement σ_{ii} konstant, so sind es auch alle übrigen σ_{ik}. M. a. W., das Spektrum kann stets bereits aus den σ_{ii} ermittelt werden. Da nun bei fastperiodischen Matrizen nach (63) alle σ_{ii} einander gleich sind, so versteht sich von selbst, daß dabei σ_{00} allein das ganze Spektrum in Evidenz setzen muß. Ebenso kann $\mathfrak{B}(f)$ nur dann ein Punktspektrum haben, wenn bereits σ_{00} einen Sprung hat.

Anmerkungen und Literatur.

Die nachstehenden Literaturangaben erheben keinen Anspruch auf Vollständigkeit, berücksichtigen vielmehr nur die wichtigeren bzw. diejenigen Arbeiten, auf welche oben Bezug genommen wurde, und wollen andererseits auf einige Gebiete hinweisen, die im Texte nur berührt werden konnten.

Im übrigen sei ein für allemal auf den unlängst erschienenen Artikel von E. Hellinger und O. Toeplitz verwiesen, Enc. der math. Wiss. II 3, 2, S. 1335 bis 1601 [auch besonders erschienen unter dem Titel Integralgleichungen und Gleichungen mit unendlich vielen Unbekannten, Leipzig 1927].

Endliche Matrizen.

Die systematische Begründung der Matrizenrechnung verdankt man G. Frobenius, Crelles Journal **84** (1878), S. 1ff. Wegen der unitären Transformationstheorie siehe I. Schur, Math. Annalen **66** (1909), S. 488ff. und O. Toeplitz, Math. Ztschr. **2** (1918), S. 187—197; vgl. auch G. Frobenius, Sitzber. d. Berliner Akad. 1911, S. 241—243. Über das Verhältnis des Spektrums zu dem Wertevorrat siehe O. Toeplitz, loc. cit. und F. Hausdorff, Math. Ztschr. **3** (1919), S. 314—316. — Die auf S. 44 (Fußnote) erwähnte Parameterdarstellung der unitären Matrizen findet man bei A. Hurwitz, Gött. Nachr. 1897, S. 82ff.; vgl. hierzu I. Schur, Sitzber. der Berliner Akad. 1924, S. 196. — Über endliche Matrizen im allgemeinen vgl. W. F. Meyer und J. Drach, Enc. des sciences math. I **2** (1911—12), S. 386ff.

Die Jacobische Transformation geht in ihren wesentlichen Zügen eigentlich auf Lagrange und Gauss zurück. Bei Jacobi tritt sie anläßlich einer Fragestellung aus der arithmetischen Theorie der quadratischen Formen auf; vgl. hierzu zum Beispiel H. Minkowski, Geometrie der Zahlen, Leipzig und Berlin 1910, S. 182ff. In diesem zahlentheoretischen Zusammenhang waren die Beziehungen zu dem Euklidischen Algorithmus von jeher bekannt; wegen der algebraischen bzw. „integralen" Theorie der Jacobischen Matrizen siehe E. Heine, Handbuch der Kugelfunktionen I, zweite Auflage, Berlin 1878, S. 415ff., sowie die weiter unten bei den unendlichen Matrizen angegebene Literatur (die formalen Zusammenhänge sind nur sehr spät, erst in den letzten zwanzig Jahren, also erst nach Stiltjes und Hilbert erkannt worden).

Die in § 28 skizzierte grundlegende Folgerung aus der Jacobischen Transformation hat zuerst wohl O. Toeplitz gezogen, Gött. Nachr. 1907, S. 101—102; vgl. auch I. Schur, Math. Ztschr. **1** (1918), S. 202—205. — Dem E. Schmidtschen Verfahren und dem bewußten Orthogonalisierungsprozesse begegnet man zuerst nicht in der Algebra, sondern in der Analysis; siehe E. Schmidt, Math. Annalen **63** (1907), S. 442, sowie die daselbst zitierte Abhandlung von J. P. Gram, ferner eine Arbeit von G. Kowalewski, Sitzber. d. Wiener Akad. **120** (1911), S. 77; vgl. auch H. Minkowskis „Adaptionsprozeß" in seinem Buche Diophantische

Approximationen, Leipzig 1907, S. 90ff.; für Vektoren siehe die weiter unten bei den unendlichen Matrizen zitierte Abhandlung von E. SCHMIDT selbst.

Zur näheren Begründung der §§ 4—7 sei auf die Lehrbücher der Algebra verwiesen [zum Beispiel E. STUDY, Einleitung in die Theorie der Invarianten linearer Transformationen auf Grund der Vektorrechnung I, Berlin 1923, oder O. PERRON, Algebra I, Berlin und Leipzig 1927]. Der Gedanke der Ausschaltung der Determinantentheorie und die Durchführung eines determinantenfreien Beweises bei den Hauptsätzen über endliche lineare Gleichungssysteme (§ 4) findet sich in der weiter unten zitierten Rendiconti-Arbeit von O. TOEPLITZ über finite Matrizen; vgl. auch eine Note von A. KNESER in der SCHWARZ-Festschrift, S. 177 (1914). Hier ist auch ein Aufsatz von R. G. D. RICHARDSON zu erwähnen, Trans. of Amer. Math. Soc. **26** (1924), S. 451—478.

In Kap. I sind Terminologie und Methodik — etwas abweichend von der herkömmlichen Darstellung [vgl. hingegen H. HAHN, Jahresbericht der deutschen Math.-Ver. **20** (1911), S. 69—117] — mehrfach im Anschluß an die FREDHOLM-HILBERT-SCHMIDTsche Theorie der Integralgleichungen getroffen worden (C. NEUMANNsche Reihen, HILBERTsche Funktionalgleichung, die Betrachtungen von § 8, usf.); vgl. freilich die eingangs zitierte klassische Arbeit von G. FROBENIUS.

Die in Kap. I durchgeführten Betrachtungen über Hermitesche Spektralmatrizen bilden nur eine algebraische Umschreibung der HILBERT-HELLINGERschen Theorie auf endliche Matrizen. Ebenso leicht ist es, eine „integrale" Spektraltheorie für alle endlichen normalen Matrizen anzugeben. Zu diesem Ende genügt es (vgl. weiter unten S. 279), die Formeln auch noch für den Spezialfall der unitären Matrizen herzuleiten, und hierfür benötigt man freilich nicht die allgemeine Theorie (§§ 98—101), man kann vielmehr ebenso wie in § 28ff. vorgehen; vgl. nämlich die Schlußbemerkung auf S. 217, und beachte, daß bei unitären Transformationen eine Hermitesche oder unitäre Matrix sich mit ihrer Spektralmatrix kogredient transformiert. — Die Formel (72), S. 51 kann in diesem Zusammenhang auch aus der Darstellungstheorie der infinitesimalen Gruppen gewonnen werden [zur Orientierung vgl. man H. WEYL, Mathematische Analyse des Raumproblems, Berlin 1923, S. 33ff.].

Die Betrachtungen von §§ 24ff. können auch wie folgt verifiziert werden. Es sei $\mathfrak{A} = \|a_{ik}\|$ eine (endliche) Hermitesche Matrix, und $\|\sigma_{ik}(\mu)\|$ ihre Spektralmatrix. Ist dann $\mathfrak{U}$ eine unitäre Matrix, so ist $\mathfrak{U}\|\sigma_{ik}(\mu)\|\mathfrak{U}^{-1}$ nach (64) und mit Rücksicht auf § 7 bzw. § 16 die Spektralmatrix von $\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}$. Wir können daher annehmen, daß $\mathfrak{A}$ eine Diagonalmatrix, also von der Gestalt $\|\lambda_i \delta_{ik}\|$ ist, wobei die n Eigenwerte λ_i nicht voneinander verschieden zu sein brauchen. Es sei $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(p)}$ das System der verschiedenen Eigenwerte und s_ν die Vielfachheit von $\lambda^{(\nu)}$, so daß $s_1 + s_2 + \ldots + s_p = n$ gilt. Als zu

$$\lambda\mathfrak{E} - \mathfrak{A} = \|(\lambda - \lambda_i)\,\delta_{ik}\|; \quad \lambda = \lambda^{(\nu)}, \quad \nu = 1, 2, \ldots, p$$

gehöriges normiert orthogonales Eigenvektorsystem kann nach S. 27 das System der n Einzelvektoren gewählt werden, woraus wegen (64) und mit Rücksicht auf § 16 die Spektralmatrix $\|\sigma_{ik}(\mu)\|$ von $\mathfrak{A} = \|\lambda_i \delta_{ik}\|$ unmittelbar abgelesen werden kann: Es gilt $\sigma_{ik}(\mu) \equiv 0$ für $-\infty < \mu < +\infty$, wenn $i \neq k$ ist, während jedes $\sigma_{ii}(\mu)$ für genau ein $\mu = \lambda^{(\nu)}$ den Sprung 1 erleidet und sonst stetig, also konstant ist, nämlich $\equiv 0$ links und $\equiv 1$ rechts von der Sprungstelle, und zwar ist die Anzahl derjenigen σ_{ii}, die den Sprung für $\mu = \lambda^{(\nu)}$ erleiden, gleich der Vielfachheit s_ν von $\lambda^{(\nu)}$. — Ebenso liegen die Verhältnisse bei den unitären und allgemeiner bei den normalen Matrizen.

Die Sonderstellung der normalen Matrizen (gegenüber den nicht notwendig normalen), insbesondere ihre Eigenschaft, daß sie die weiteste Klasse darstellen, innerhalb deren eine *unitäre* Spektraltheorie überhaupt möglich ist (so daß das Spektrum das *vollständige* Invariantensystem der unitären Transformationsgruppe bildet; vg. S. 25ff.), klärt sich eigentlich nur dann auf, wenn man zu der Invariantentheorie beliebiger, also nicht notwendig normaler Matrizen, d. h. zu der Theorie der Elementarteiler übergeht, wobei dann der Natur der Sache nach auch nichtunitäre Transformationen zugelassen werden müssen. Doch konnte in diesem Buch auf diese von WEIERSTRASS herrührende Lehre nicht eingegangen werden. Moderne Darstellungen dieser etwas schwerfälligen Theorie findet man bei WEYL, loc. cit., sowie in dem von W. KRULL verfaßten Anhange des zweiten Bandes von O. HAUPTS Einführung in die Algebra, Leipzig 1929, S. 621ff.; vgl. W. KRULL, Sitzber. der Heidelberger Akad. 1926, Nr. 1. — Vom Standpunkte der Elementarteilertheorie aus muß es eigentlich als ein „Zufall" bezeichnet werden, daß die Normalspur

$$\vartheta(\mathfrak{A}) = \sum_{i=1}^{n} \sum_{k=1}^{n} |a_{ik}|^2$$

nach S. 28 auch bei nicht normalen $\mathfrak{A}$ unitär invariant ist: $\vartheta(\mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1}) = \vartheta(\mathfrak{A})$, auch wenn $\mathfrak{A}\mathfrak{A}^* \neq \mathfrak{A}^*\mathfrak{A}$, wenn nur $\mathfrak{U}^{-1} = \mathfrak{U}^*$.

Wie FROBENIUS in seiner zweiterwähnten Arbeit bemerkt hat, kann $\vartheta(\mathfrak{A})$ zur „Metrisation" der Mannigfaltigkeit der n-ären Matrizen verwendet werden. Legt man nämlich die „Entfernung" $\Delta(\mathfrak{A}, \mathfrak{B})$ zweier Matrizen $\mathfrak{A}, \mathfrak{B}$ durch

$$\Delta(\mathfrak{A}, \mathfrak{B}) = \vartheta(\mathfrak{A} - \mathfrak{B})$$

fest, so ist offenbar $\Delta(\mathfrak{A}, \mathfrak{B}) = \Delta(\mathfrak{B}, \mathfrak{A})$, ferner $\Delta(\mathfrak{A}, \mathfrak{B}) \geqq 0$, wobei das Gleichheitszeichen dann und nur dann gilt, wenn die zwei Matrizen zusammenfallen, endlich ist, wie der Leser selbst verifizieren möge, auch das sogenannte Dreiecksaxiom erfüllt:

$$\Delta(\mathfrak{A}, \mathfrak{B}) \leqq \Delta(\mathfrak{A}, \mathfrak{C}) + \Delta(\mathfrak{C}, \mathfrak{B}).$$

Bei den in Betracht kommenden unendlichen Matrizen steht dieses wertvolle Hilfsmittel der Darstellungstheorie leider nicht mehr zur Verfügung, da die Reihe ϑ bereits für die unendliche Einheitsmatrix divergiert.

Literatur zu S. 46.: O. HÖLDER, Gött. Nachr. 1889, S. 38—47; vgl. auch J. L. W. V. JENSEN, Acta Math. **30** (1906), S. 175ff. — E. HELLY, Monatshefte f. Math. **31** (1921), S. 60—91. — H. BOHR, Gött. Nachr. 1913, S. 411ff. — O. TOEPLITZ, Gött. Nachr. 1913, S. 417—432. — M. RIESZ, Acta Math. **49** (1926), S. 465ff. — Auch die Theorie der Laurentmatrizen sowie ihrer Analogien und Verallgemeinerung ist hier zu erwähnen.

Stieltjessche Integrationstheorie u. dgl.

Die STIELTJESsche Integrationstheorie trat zuerst im Rahmen der Kettenbruchlehre auf (vgl. S. 241), siehe TH. J. STIELTJES, Ann. de Toulouse **8** (1894) [J], S. 1–122; **9** (1895) [A], S. 1–47 oder Œuvres complètes II, Groningen 1918, S. 402–566. Unter den Lehrbüchern sind zu empfehlen O. PERRON, Die Lehre von den Kettenbrüchen, Leipzig und Berlin 1913, insbes. Kap. IX, sowie H. LEBESGUE, Leçons sur l'intégration et la recherche des fonctions primitives. Deuxième édition. Paris 1928 (Chap. XI), ferner die auf S. 75 zitierte Schrift von W. BLASCHKE. Wegen Fragen der reellen Funktionentheorie sei ein für allemal verwiesen auf die von A. ROSENTHAL bearbeiteten vorzüglichen Artikel in der Enc. der math.

Wiss. II 3, 2 (1923), sowie an die bekannten Bücher von C. CARATHÉODORY und H. HAHN. Eine gute Einführung in die LEBESGUEsche Theorie gibt E. KAMKE, Das LEBESGUEsche Integral. Eine Einführung in die moderne Theorie der reellen Funktionen. Leipzig und Berlin, 1925. — Ich mußte mich aus äußeren Gründen leider entschließen, den BOREL-LEBESGUEschen Gedankenkreis nicht grundsätzlich als bekannt vorauszusetzen. Ganz konsequent ließ sich dies freilich nicht durchführen, und so ist es an manchen Stellen zu einer Art Kompromiß gekommen.

Die STIELTJESsche Umkehrformel findet sich in der großen Abhandlung von STIELTJES, tritt aber in einer etwas versteckten Form und ohne Zugrundelegung des allgemeinen STIELTJESschen Integralbegriffes bereits bei CH. HERMITE auf [vgl. z. B. Crelles Journal **91** (1881), S. 54ff.]. Die HILBERTsche Residuenformel ist offenbar weniger scharf als die STIELTJESsche; sie liefert ja nur die Sprünge der Belegung, und kann zwischen zwei stetigen Belegungen keinen Unterschied machen. Wegen der auf S. 100 erwähnten paradox klingenden Erscheinungen siehe insbesondere A. PRINGSHEIM, Münchener Sitzber. 1927, S. 145ff., sowie gewisse Untersuchungen von GOURSAT, BOREL, POMPÉIU, STEFFENSEN, HAAR, DENJOY, CARLEMAN und anderen. — C. NEUMANN hat mittels der zugeordneten Kugelfunktionen für alle Funktionen (genauer Funktionenzweige) $F(\lambda)$, die sich auf der längs des Intervalles $-1 \leq \lambda \leq +1$ aufgeschlitzten λ-Ebene mit Einschluß des unendlich fernen Punktes regulär verhalten, eine Darstellung der Form

$$F(\lambda) = F(\infty) + \sum_{n=1}^{\infty} \int_{-1}^{+1} \frac{d\,\sigma_n(\mu)}{(\lambda-\mu)^n}$$

gegeben, wobei die σ_n sogar stetig differentiierbar sind. Doch handelt es sich dabei, im Gegensatz zu (152), um keine eindeutige „kanonische" Darstellung von $F(\lambda)$, d. h. es sind „Nullentwicklungen" möglich, bei welcher nicht alle σ_n konstant sind, trotzdem die Reihe (so gut wie nur möglich) gegen die Funktion $F(\lambda) \equiv 0$ konvergiert.

Der wesentliche Inhalt der HILBERTschen Residuenformel besagt, daß, sobald die Belegung $\sigma(\mu)$ von beschränkter Schwankung ist, der Grenzwert

$$\lim_{\varepsilon=0} \varepsilon \int_{-\infty}^{+\infty} \frac{d\sigma(\nu)}{\mu+\varepsilon-\nu}; \qquad \varepsilon > 0$$

für alle reellen μ existiert und dann und nur dann verschwindet, wenn μ eine Stetigkeitsstelle von σ ist. Dieser Satz ist mit unwesentlichen Einschränkungen unlängst von I. SCHOENBERG wiedergefunden und im Falle einer durchweg stetigen Funktion mittels einer LAPLACEschen Transformation auf eine mehr handliche Gestalt gebracht worden; siehe Math. Ztschr. **28** (1928), S. 179ff.; vgl. auch **31** (1929) [unter der Presse]. Den auf S. 100 erwähnten Zusammenhang der STIELTJESschen Umkehrformel mit der FOURIERschen hat man mehrfach bemerkt [vgl. insbesondere G. H. HARDY, Messenger of Math. **54** (1924), S. 22], ohne bisher ihn näher verfolgt zu haben. Man erkennt am besten, worum es sich bei diesem Zusammenhange handelt, wenn man die Herleitung der Formeln (72) oder (316) [vgl. A. WINTNER, Annalen der Physik **81** (1926), S. 577—586, 846—854] mit den Methoden der Theorie der LAPLACEschen Transformation vergleicht [zur Orientierung hierüber vgl. den Bericht von G. DOETSCH, Jahresbericht der deutschen Math.-Ver. **36** (1927), S. 19—30]. Zur FOURIERschen Umkehrformel unter Zugrundelegung des STIELTJESschen Integralbegriffes vgl. z. B. H. HAHN, Acta Math.

49 (1926), S. 301—353; N. WIENER, Math. Ztschr. 24 (1925), S. 575ff.; N. WIENER, Proc. London Math. Soc. 27 (1927), S. 483ff., über die auf S. 100 erwähnten CAUCHYschen Hauptwerte außer älteren Arbeiten (vgl. die zitierte Note von HARDY) ein Ergebnis von S. POLLARD und Miss R. C. YOUNG, Proc. London Math. Soc. 28 (1928), S. 293ff. bzw. S. POLLARD, Journ. London Math. Soc. 2 (1927), S. 37ff.; vgl. übrigens G. HERGLOTZ, Sächsische Sitzungsberichte 1923, S. 31ff., endlich über die sogenannten singulären Integralen im allgemeinen H. LEBESGUE, Ann. de Toulouse (3) 1 (1909), S. 25—117. Die für die Spektraltheorie der normalen Matrizen (vgl. S. 203 und S. 280) grundlegenden Untersuchungen von J. RADON sind in den Sitzber. der Wiener Akad. erschienen, 122 (1913), S. 1295—1433, vgl. 128 (1919), S. 1083ff. — Wegen der Zurückführung der STIELTJESschen Theorie auf die LEBESGUEsche siehe die Note von H. LEBESGUE, C. R. 150 (1910), S. 86—88, welche den Untersuchungen von von H. HAHN, Monatshefte für Math. 23 (1912), S. 161—224 zum Ausgangspunkt diente.

Im übrigen sowie wegen der ganz neuen Literatur siehe den bereits erwähnten Bericht von A. ROSENTHAL.

Die neuzeitlichen Untersuchungen, auf welche in §§ 34—49 Bezug genommen wurde, sind die folgenden: E. HELLY, Sitzber. der Wiener Akad. 121 (1912), S. 265 bis 297; J. GROMMER, Crelles Journal 144 (1914), S. 114ff.; E. HELLINGER und O. TOEPLITZ, ibid. S. 212ff.; H. HAMBURGER, Math. Ann. 81 (1920), S. 266ff.; F. RIESZ, Ann. de l'École Norm. Sup. 28 (1911), S. 33—62 und 31 (1914), S. 9—14. Über die Vertauschung der Reihenfolge zweier Integrationen (§ 42) vgl. L. LICHTENSTEIN, Gött. Nachr. 1910, S. 474—475, über die Stieltjesche Integration unstetiger Funktionen (§ 88) R. SCHMIDT, Math. Ztschr. 22 (1925), S. 123—125. — Über Auswahlsätze in abstrakten Räumen u. dgl. (S. 82) vgl. F. HAUSDORFF, Mengenlehre, Zweite Auflage, Berlin und Leipzig 1927, Kap. VI, insbesondere bei konvexen Körpern W. BLASCHKE, l. c., bei konvexen Funktionen einer Veränderlichen (deren Theorie mit derjenigen der monotonen Funktionen aufs engste zusammenhängt) W. BLASCHKE und G. PICK, Math. Ann. 77 (1916), S. 277—300, endlich in der komplexen Funktionentheorie (VITALI-MONTEL) L. BIEBERBACH, Enc. der math. Wiss. II 3, 1 (1921), S. 491ff.

Wegen § 50 vgl. die Göttinger Dissertation (1907) von E. HELLINGER, Über die Orthogonalinvarianten quadratischer Formen von unendlich vielen Veränderlichen, sowie seine Habilitationsschrift, Crelles Journal 136 (1909), S. 210—271.

Unendliche Matrizen.

Wegen der Entstehung der Theorie der unendlich vielen Variablen siehe den Enzyklopädieartikel von E. HELLINGER und O. TOEPLITZ; vgl. auch eine historische Bemerkung bei A. WINTNER, Ztschr. für Phys. 48 (1928), S. 509.

Die grundlegende Arbeit von D. HILBERT ist seine Vierte Mitteilung über Integralgleichungen, Gött. Nachr. 1906, S. 157—227. Wegen Konvergenzfragen bei beschränkten Bilimarformen u. dgl. siehe E. HELLINGER und O. TOEPLITZ, Math. Annalen 69 (1910), S. 289—330; F. RIESZ, Math. Annalen 69 (1910), S. 449ff.; I. SCHUR, Crelles Journal 140 (1910), S. 1—28; O. TOEPLITZ, Gött. Nachr. 1910, S. 503, wegen der Weiterführung der HILBERTschen Theorie der beschränkten Hermiteschen (eigentlich: reell-symmetrischen) Matrizen die Dissertation von HELLINGER, wegen der unitären Matrizen und wegen spektralen Fragen bei beschränkten nicht Hermiteschen Matrizen A. WINTNER, Math. Ztschr. 30 (1929), S. 228—282, wegen des Reziprokenproblems O. TOEPLITZ, Gött. Nachr. 1907, S. 101—109 und E. HILB, Erlanger Sitzungsberichte 40 (1908), S. 84—85, wegen der Ergänzung von nicht

vollständig unitären Matrizen zu unitären (S. 158ff.) D. HILBERT, loc. cit. S. 196, E. SCHMIDT, loc. cit. S. 61ff., F. RIESZ, Gött. Nachr. 1907, S. 120, wegen unendlicher nichtlinearer Systeme von Gleichungen (und Differentialgleichungen) unter Zugrundelegung einer dem nichtlinearen Falle angepaßten Verallgemeinerung des HILBERTschen Beschränktheitsbegriffes A. WINTNER, Math. Ztschr. **28** (1928), S. 451—470. — Eine interessante Darstellung der HILBERTschen Theorie gab F. RIESZ, Les systémes d'équations linéaires à une infinité d'inconnues. Paris 1913 (diese Monographie war die erste zusammenfassende Darstellung der Theorie, und ist sehr zu empfehlen); vgl. ferner F. RIESZ, Gött. Nachr. 1910, S. 190—195.

Über nicht beschränkte Hermitesche Matrizen siehe D. HILBERT, loc. cit., S. 164ff., über Matrizen, bei welchen die Reihen

$$\sum_{k=1}^{\infty} |a_{ik}|^2, \qquad i = 1, 2, \ldots$$

konvergent sind, E. SCHMIDT, Rend. Palermo **25** (1928), S. 53—77 und E. HILB, Math. Annalen **70** (1911), S. 79—86, über finite Matrizen O. TOEPLITZ, Rend. Palermo **28** (1909), S. 88—96, H. BOHR, Danske Wid. Selsk. Medd. **7** (1925), Nr. 1, über halbbeschränkte Matrizen A. WINTNER, Sächsische Sitzungsberichte **79** (1927), S. 145—157, über die CARLEMANsche Theorie T. CARLEMAN, Sur les équations integrales singulières à noyau reél et symétrique, Uppsala Univers. Årsskrift, Uppsala 1923. Eine Weiterführung der CARLEMANschen Theorie hat vor einiger Zeit J. v. NEUMANN in Aussicht gestellt, Jahresbericht der deutschen Math.-Ver. **37** (1928), S. *11*—*14* [die ausführliche Darstellung wird demnächst in den Math. Annalen erscheinen].

Über JACOBIsche Matrizen, assoziierte Kettenbrüche und Momentenproblem siehe O. TOEPLITZ, Gött. Nachr. 1910, S 489ff.; J. GROMMER, Crelles Journal **144** (1914), S. 114—166; E. HELLINGER und O. TOEPLITZ, ibid., S. 212—238 und S. 318; H. HAMBURGER, Math. Ann. **81** (1920), S. 235–319, **82** (1921), S. 120–164, S. 168–187; E. HELLINGER, Math. Annalen **86** (1922), S. 18—29; T. CARLEMAN, loc. cit.; M. RIESZ, Arkiv för Mat., Astr. och Fysik **17** (1913), Nr. 16; A. WINTNER, Math. Ztschr. **30** (1929), S. 285—289; die älteren Untersuchungen findet man in dem auf S. 275 zitierten Buche von O. PERRON; vgl. auch G. POLYA und G. SEGÖ, Aufgaben und Lehrsätze der Analysis I, Berlin 1925, S. 286 und 314.

Über die TOEPLITZschen und LAURENTschen Matrizen sowie das trigonometrische Momentproblem vgl. O. TOEPLITZ, Rend. Palermo **32** (1911), S. 191—192, Math. Annalen **70** (1911), S. 351—376, Math. Ztschr. **12** (1922), S. 189—200; I. SCHUR, Crelles Journal **147** (1917), S. 205—232; A. WINTNER, Math. Ztschr. **30** (1929), S. 277—281; ferner F. RIESZ, Ann. de l'École Normale Sup. **28** (1911), S. 54ff.; G. HERGLOTZ, Sächs. Sitzungsber. **63** (1911), S. 501—511; F. HAUSDORFF, Math. Ztschr. **16** (1923), S. 222ff.

Wegen des Zusammenzuges der vollständigen orthogonalen Funktionssysteme mit der Spektraltheorie siehe D. HILBERT, loc. cit., S. 206ff.; E. HELLINGER, Dissertation, S. 61ff.; E. HELLINGER und O. TOEPLITZ, Crelles Journal **144** (1914), 212ff. — Die Einordnung der Theorie der Spektralmatrix in diejenige der Gruppendarstellungen liegt bisher in der Literatur nicht vor; da dies im Rahmen dieser Schrift wegen Raummangel nicht möglich war, muß hierüber auf eine demnächst in der Math. Ztschr. erscheinende Untersuchung verwiesen werden. Vgl. übrigens die neuesten Ergebnisse über kontinuierliche Gruppen, insbes. F. PETER und H. WEYL, Math. Ann. **97** (1927), S. 737—755.

Zu § 120 vgl. D. HILBERT, loc. cit.; O. TOEPLITZ, Gött. Nachr. 1913, 432; H. WEYL, Rend. Palermo **27** (1909), S. 373—392; T. CARLEMAN, loc. cit.

Über fastperiodische Matrizen siehe W. HEISENBERG, Math. Ann. **95** (1926), S. 685—688 und A. WINTNER, Math. Ztschr. **30** (1929), S. 290—319, über fastperiodische Funktionen H. BOHR, Acta Math. **45** (1924), S. 29—121, **46** (1925), S. 101—204, **47** (1926), S. 237—281 und H. WEYL, Math. Annalen **97** (1927), S. 338—356.

Über die sich auf unendliche Matrizen beziehenden mathematischen Probleme der Quantenmechanik kann sich der Mathematiker am bequemsten orientieren aus der vorher erwähnten Arbeit von HEISENBERG sowie aus dem Büchlein von M. BORN, Probleme der Atomdynamik, Berlin 1926.

Radonsche Integrale und normale Spektraltheorie.

Nach Beendigung der Drucklegung habe ich bemerkt, daß auf die Hermitesche und auf die unitäre Spektraltheorie eine solche aller normalen Matrizen gegründet werden kann.

Zunächst kann man nämlich zu jeder beschränkten normalen Matrix $\mathfrak{A}$, für welche $\mathfrak{A}^{-1}$ existiert, auf genau eine Weise eine positiv definite Matrix $\mathfrak{P}$ und eine unitäre Matrix $\mathfrak{U}$ derart angeben, daß $\mathfrak{A} = \mathfrak{P}\mathfrak{U} = \mathfrak{U}\mathfrak{P}$ wird (während umgekehrt jede Matrix $\mathfrak{A}$, die auf diese Weise zerlegbar ist, normal ist und eine Reziproke hat). Aus leicht ersichtlichen Gründen soll $\mathfrak{P}$ als der Absolutbetrag, $\mathfrak{U}$ als das Azimut von $\mathfrak{A}$ bezeichnet werden. Ist z. B. $\mathfrak{A}$ die fastperiodische Matrix der fastperiodischen Funktion $f(t)$, die der Null nicht beliebig nahe kommen kann, so daß $\mathfrak{A}^{-1}$ existiert, so ist die fastperiodische Matrix von $|f(t)|$ der Absolutbetrag, diejenige von $f(t)\,|f(t)|^{-1}$ das Azimut von $\mathfrak{A}$; vgl. die Schlußbemerkung meines zuletzt erwähnten Aufsatzes, woselbst (§ 9) nicht diese, sondern eine kompliziertere, „kartesische“ Zerlegung angedeutet wird (siehe weiter unten).

Es sei nun $\|\sigma_{pq}(\mu)\|$ die Spektralmatrix des Absolutbetrages, $\|\tau_{pq}(\varphi)\|$ diejenige des Azimuts der normalen Matrix $\mathfrak{A}$. Da die beiden Matrizen $\mathfrak{P}$ und $\mathfrak{U}$ vertauschbar sind, so sind es auch ihre Resolventen, und daher, wie die sinngemäße Wiederholung einer Hilbertschen Schlußweise zeigt, auch ihre Spektralmatrizen $\|\sigma_{pq}\|$, $\|\tau_{pq}\|$, und zwar bei jedem Wertepaare (μ, φ). Da diese beiden Spektralmatrizen Hermitisch sind, so ist es mithin auch ihr Produkt

$$\|\zeta_{pq}(\mu, \varphi)\| = \|\sigma_{pq}(\mu)\| \cdot \|\tau_{pq}(\varphi)\|; \quad -\infty < \mu < +\infty, \quad 0 \leqq \varphi \leqq 2\pi.$$

Die Matrix $\|\zeta_{pq}\|$ ist nun als die Spektralmatrix der normalen Matrix $\mathfrak{A}$ anzusprechen. Es zeigt sich nämlich, daß für hinreichend große $|\lambda|$

$$(\lambda\mathfrak{E} - \mathfrak{A})^{-1} = \left\| \int_{-\infty}^{+\infty}\int_{0}^{2\pi} \frac{d\zeta_{pq}(\mu, \varphi)}{\lambda - \mu e^{i\varphi}} \right\|$$

gilt, wobei das Integral, wie bereits aus der Schreibweise hervorgeht, nicht als wiederholtes Stieltjessches Integral, sondern als *Radonsches Doppelintegral* zu verstehen ist [die Radonsche Monotonieeigenschaft der Kopplungsform von $\|\zeta_{pq}(\mu, \varphi)\|$ kann aus bekannten Eigenschaften der klassischeu Spektralmatrix leicht geschlossen werden]. Da $\mathfrak{P}$ beschränkt ist, so gilt für hinreichend große $|\mu|$, wie wir wissen, $d\sigma_{pq}(\mu) \equiv 0$, und daher *in dem Radonschen Sinne*

$d\zeta_{pq}(\mu, \varphi) \equiv 0$, so daß der Integrationsbereich ganz im Endlichen liegt. Da ferner $\mathfrak{P}$ positiv definit ist, so gibt es ein $\varepsilon > 0$ derart, daß $d\sigma_{pq}(\mu) \equiv d\zeta_{pq}(\mu, \varphi) \equiv 0$ wird für $-\infty < \mu < \varepsilon$. Wir können daher, wenn wir noch den Buchstaben μ auf ϱ vertauschen, einfach

$$(*) \qquad (\lambda\mathfrak{E} - \mathfrak{A})^{-1} = \left\| \int\limits_0^{+\infty} \int\limits_0^{2\pi} \frac{d\zeta_{pq}(\varrho, \varphi)}{\lambda - \varrho e^{i\varphi}} \right\|; \qquad \zeta_{pq}(\varrho, \varphi) = \sum_{j=1}^{\infty} \sigma_{pj}(\varrho)\, \tau_{jq}(\varphi)$$

setzen. Es handelt sich also bei der Zerlegung $\mathfrak{A} = \mathfrak{P}\mathfrak{U}$ tatsächlich um die „Polarkoordinaten" von $\mathfrak{A}$. Bei der „kartesischen" Zerlegung $\mathfrak{A} = \mathfrak{G} + i\mathfrak{H}$ in Hermitesche Komponenten baut sich die Spektralmatrix von $\mathfrak{A}$ aus denjenigen von $\mathfrak{G}$ und $\mathfrak{H}$ auf eine viel kompliziertere Weise auf (was davon herrührt, daß die kartesische Zerlegung additiv und nicht multiplikativ ist). Die Existenz einer solchen kartesischen Integraldarstellung der Resolvente *aller* beschränkten normalen Matrizen haben von Neumann und ich in während der zweiten Hälfte vom Januar d. J. stattgefundenen Gesprächen mit Bestimmtheit vermutet. Die Existenz dieser Formel kann jetzt auf Grund von (*) sichergestellt werden, doch ist (*) viel bequemer, so daß ich nunmehr auch bei den fastperiodischen Matrizen die Darstellung (*) bevorzugen werde.

Es war bisher vorausgesetzt, daß $\mathfrak{A}^{-1}$ existiert. Ist dies nicht der Fall, so wähle man eine reelle Zahl ν derart, daß die (offenbar normale) Matrix $\mathfrak{A} + \nu\mathfrak{E}$ eine Reziproke hat, und bezeichne durch $\| \zeta_{pq}^{(\nu)}(\varrho, \varphi) \|$ die Spektralmatrix von $\mathfrak{A} + \nu\mathfrak{E}$. Dann hängt die Matrix $\| \zeta_{pq}^{(\nu)}(\varrho - \nu, \varphi) \|$ von ν gar nicht ab, und bezeichnet man sie durch $\| \zeta_{pq}(\varrho, \varphi) \|$, so bleibt die Integraldarstellung (*) bestehen.

Das Spektrum von $\mathfrak{A}$ besteht jetzt nicht mehr aus allen Nichtkonstanzstellen der ζ_{pq} und deren Häufungsstellen, das Punktspektrum entsprechend nicht aus allen, sondern nur aus denjenigen Unstetigkeitsstellen, in welchen „$d\zeta_{pq}(\varrho, \varphi) \neq 0$" gilt bei mindestens einem Wertepaare (p, q). — Nach Radon sind sogar alle Unstetigkeitsstellen auf einer Halbstrahlenfolge $\varphi = \varphi_m$; $m = 1, 2, \ldots$ und auf einer Kreislinienfolge $\varrho = \varrho_n$; $n = 1, 2, \ldots$ gelegen.